A Corpus-based Register Analysis of Corporate Blogs – Text Types and Linguistic Features

Thesis submitted in accordance with the requirements of
The University of Liverpool for the degree of Doctor in Philosophy
by
Yang Wu

August 2015

ACKNOWLEDGEMENTS

The writing of this thesis would not have been finished without the support of numerous people. First of all, I would like to express my sincere gratitude to my supervisors, Dr. Zhoulin Ruan, Professor Michael Hoey, and Dr. Songqing Li, who have provided me with great support throughout the current study. Without their insightful advice, guidance, encouragements, and unselfish help, I could never finish my doctoral study in Xi'an Jiaotong-Liverpool University (XJLTU). For me, my supervisors are not only teachers, but also my lifetime friends.

My sincere thanks also go to the member of the Independent Assessment Panel, Dr. Paul Cheung and Dr. Bin Zou, for their sharing of valuable suggestions and experiences with me. The staff in the postgraduate school of XJTLU and my colleagues and peers in Shanghai University are thanked for their help and support.

I am very grateful to Mr. Bai Wei and Mr. Wu Jingjin, for their help on the manipulation of python running on UNIX-like systems. My gratitude also goes to Mr. Xiang Hongyong and Mr. Liu Junshu who help me with much information on how to set up and use EndNote.

Finally, I would like to thank my parents, my husband, and my child for their love, support, and patience over the last few years.

ABSTRACT

A main theme in sociolinguistics is register variation, a situation and use dependent variation of language. Numerous studies have provided evidence of linguistic variation across situations of use in English. However, very little attention has been paid to the language of corporate blogs (CBs), which is often seen as an emerging genre of computer-mediated communication (CMC). Previous studies on blogs and corporate blogs have provided important information about their linguistic features as well as functions; however, our understanding of the linguistic variation in corporate blogs remains limited in particular ways, because many of these previous studies have focused on individual linguistic features, rather than how features interact and what the possible relations between forms (linguistic features) and functions are. Given these limitations, it would be necessary to have a more systematic perspective on linguistic variation in corporate blogs.

In order to study register variation in corporate blogs more systematically, a combined framework rooted in Systemic Functional Linguistics (SFL), and register theories (e.g., Biber, 1988, 1995; Halliday & Hasan, 1989) is adopted. This combination is based on some common grounds they share, which concern the functional view of language, co-occurrence patterns of linguistic features, and the importance of large corpora to linguistic research. Guided by this framework, this thesis aims to: 1) investigate the functional linguistic variations in corporate blogs, and identify the text types that are distinguished linguistically, as well as how the CB text types cut across CB industry-categories, and 2) to identify salient linguistic differences across text types in corporate blogs in the configuration of the three components of the context of situation - field, tenor, and mode of discourse. In order to achieve these goals, a 590,520-word corpus consisting of 1,020 textual posts from 41 top-ranked corporate blogs is created and mapped onto the combined framework which consists of Biber's multi-dimensional (MD) approach and Halliday's SFL. Accordingly, two sets of empirical analyses are conducted one after another in this research project. At first, by using a corpus-based MD approach which applies multivariate statistical techniques (including factor analysis and cluster analysis) to the investigation of register variation, CB text types are identified; and then, some linguistic

features, including the most common verbs and their process types, personal pronouns, modals, lexical density, and grammatical complexity, are selected from language metafunctions of mode, tenor and field within the SFL framework, and their linguistic differences across different text types are analysed. The results of these analyses not only show that the corporate blog is a hybrid genre, representing a combination of various text types, which serve to achieve different communicative purposes and functional goals, but also exhibit a close relationship between certain text types and particular industries, which means the CB texts categorized into a certain text type are mainly from a particular industry. On this basis, the lexical and grammatical features (i.e., the most common verbs, pronouns, modal verbs, lexical density and grammatical complexity) associated with Halliday's metafunctions are further explored and compared across six text types. It is found that language features which are related to field, tenor and mode in corporate blogs demonstrate a dynamic nature: centring on an interpersonal function, the online blogs in a business setting are basically used for the purposes of sales, customer relationship management and branding.

This research project contributes to the existing field of knowledge in the following ways: Firstly, it develops the methodology used in corpus investigation of language variation, and paves the way for further research into corporate blogs and other forms of electronic communication and, more generally, for researchers engaging in corpus-based investigations of other language varieties. Secondly, it adds greatly to a description of corporate blog as a language variety in its own right, which includes different text types identified in CB discourse, and some linguistic features realized in the context of situation. This highlights the fact that corporate blogs cannot be regarded as a simple discourse; rather, they vary according to text types and context of situation.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1: Introduction

1.1 Background

Variability in language is a common phenomenon, since “no normal human being talks the same way all the time” (Hymes, 1984, p. 44). In daily language, the word *style* is often used to cover different types of varieties, e.g., formal style, slang language style, business styles; however, linguists have been inclined to explore social variation according to both the uses and users of language (Halliday, McIntosh, & Strevens, 1964). Examples of user-dependent variation include dialect variation, gender variation, etc. In contrast, the use-dependent variation of language is possibly equivalent to *register*, which is more specifically referred to as “the clustering of semantic features according to situation type” (Halliday, 1978, p. 68), or “a configuration of semantic resources that the member of a culture typically associates with a situation type” (Halliday, 1978, p. 111). The current research focuses on *register*, the variation of language “according to use” (Halliday et al., 1964, p. 77). From the description and theorization of registers and register variation commencing in the 1960s (e.g., Ghadessy, 1993; Halliday et al., 1964; Steiner, 2004; Ure, 1969) to some more recent empirical register studies (e.g., Biber, 1995, 2006; Biber & Kurjian, 2006; Grieve, Biber, Friginal, & Nekrasova, 2011), new insights into the nature and function of different language phenomena have been provided consistently. This fundamentally relies on the emergence and availability of various language resources, particularly those by using communication tools through Computer-Mediated Communication (CMC) (e.g., Anis, 2007; Barnes, 2003; Crystal, 2001, 2008; Herring, 2001; Herring, 2007). CMC is a rather broad category referring to any internet-based media types that are in essence human-to-human communications via computer networks (December, 1996), such as email and chat in the first generation, wikis and blogs in the second generation, and podcasting and gaming in the third generation (Thorne, 2008). There is commonly a two-way distinction between CMC types concerning whether it is text-based or audio/video-based, and whether it operates in real time or not (Luppicini, 2007; Nguyen, 2008). On the one hand, via some CMC types, such as chat rooms, or instant

messengers, all participants are online and exchange opinions at the same time; while the interaction in some other CMC modes, such as e-mail, web blog, newsgroups, etc., does not need to be simultaneous, which allows participants more time to read, and respond to the written messages. On the other hand, in terms of the media channels of CMC, most CMC currently in use is text-based, rather than audio and video-based. The characteristics of different modes of CMC vary depending on the kind of messaging system used and the social and cultural context (Herring, 2001), which have important implications for the interpretation of the computer-mediated language. Research on CMC is typically based on text forms of CMC, and characterizes not only linguistic features, but structural features. Within studies on linguistic features of CMC, researchers have attempted to analyse the distribution of lexical and grammatical features and compare the results with features associated with written and spoken language. Earlier researchers (Ferrara, Brunner, & Whittemore, 1991; Murray, 1988) have generalized computer-mediated language as a single genre, considering that CMC is either a type of written or spoken language, or sharing features of both. More recently, people become aware of the fact that CMC is a rather diverse and complex genre influenced by various linguistic and extra-linguistic factors (Baron, 1998; Hård af Segerstad, 2002). It becomes obvious that CMC, to different extent, corresponds to both sides of the "speaking/writing divide" (Crystal, 2001, p. 28), but could not be simply put in one of those categories (written and spoken language) (Baron, 1998; Yates, 1996).

Overall, as an umbrella term encompassing various media types (e.g., online chat, email, blogs, twitters, etc.), CMC is assumed to be related to the particular medium (Kern, 2006) and shaped by sets of communicative purposes (e.g., Martin, 1984; Swales, 1990). In the domain of organizational communications, great importance has been placed on the use of different CMC types, e.g., blogs, Facebook, and Twitter (micro-blog), which are considered to be an inexpensive and "user-friendly" means of "sharing user-generated material" (Fischer & Reuber, 2011, p. 2). All these media types have their own characteristics and specialities. In terms of the communications on Twitter and Facebook, they are different in their levels of proximity. Twitter, unlike Facebook, allows users to be anonymous; and thus, it is argued that information on Twitter may be perceived as less trustworthy (Kiousis, 2001; Tsfati, 2003). If comparing

blogs and micro-blogs, the microblog differs from a traditional blog in that its content is typically smaller (140 character) which may influence its corporate engagement, while blogs are strong in content control as there are no restrictions to create and maintain the amount of posts (Yoo & de Zúñiga, 2014). What will be explored in this research project is the corporate blog (CB), which is a particular CMC genre and also a sub-type of blogs. Presented below is a brief overview of both blogs and corporate blogs.

The use of blogs is relatively a new phenomenon in the use of Internet and CMC. As a form of CMC, the origin of blogs could be traced back to 1990s. The term *blog* is an abbreviated form of *weblog*, which was coined by Jorn Barger in 1997 (Grieve et al., 2011; Kaiser & Müller-Seitz, 2008). At first, the blog was used by those personal bloggers/blog writers to publish online journals (Quible, 2005). Later, it became easier for ordinary people, rather than those technical staff to blog for personal diaries, reports, as well as all sorts of other uses since the production of tools by two small software companies, Pitas and Pyra, in 1999. Ever since the first launch of a blog between 1994 and 1998 (Gurak, Antonijevic, Johnson, Ratliff, & Reyman, 2004), there is a consistent rise and evolution of the Blogosphere, as can be seen in Technorati[1]'s State of the Blogosphere. It is particularly after 2004 that there is an explosive growth and maturing of this new media. Myers (2010) commented that those most widely read bloggers started between 2001 and 2005, and there are still lots of people starting blogging every year. Simultaneously, blogging was gradually becoming a commonly used tool in different disciplines, such as politics, psychology, etc. This gives rise to the classification of different blog sub-types, such as politics blogs for people involved in politics and policy developments, law blogs for legal marketing and business development, psychology blogs providing resources on psychological issues applied to business, work, individuals and organizations, etc. (Cowen, 2004). In any sub-types of blogs, the technical expertise in setting up a blog is not required; rather, what the blog writers need to do is using a blog. For those blog writers specializing in a particular field, they can use their blogs to publish and distribute their acquired

[1] Technorati is a publisher-advertising platform that serves as an advertising solution for the thousands of websites in its network. Technorati launched its ad network in 2008, and is currently one of the largest ad networks boasting more than 100 million unique visitors per month. The name Technorati is a blend of the words technology and literati, which invokes the notion of technological intelligence or intellectualism. (from Wikipedia)

knowledge about the subject area mostly in a conversational voice (Brady, 2005; Weil, 2006). Besides, the blog is characterized by some commonly shared structural and technological features, e.g., dated entries displayed in reverse chronological sequence, frequent updating, etc. (Bortree, 2005; Buckingham & Willett, 2013; Kelleher & Miller, 2006; Kwasnik et al., 2005; Schmidt, 2007; Weil, 2006). These features also characterize corporate blogs, the blog used in an institutional context to reach organizational goals (e.g., Sifry, 2004; Weil, 2006). For any firms delivering a blog, they have particular strategic goals which could be selling products or services, building customer loyalty, and strengthening the brand of the organization (e.g., Potts, 2007; Singh & Singh, 2008). The keys to a successful blog mainly include posting useful content and providing an interface that is easy for blog readers to navigate (Singh & Singh, 2008). If the interface of CB is easy to navigate, it can be hypothesized that the content plays an important role in blog's success. What is presented below is a brief introduction to the content and structural features of CB, which lays foundation for the creation of a corpus in this research. Usually, blogs tend to have a few things in common, as can be seen in Figure 1.1 and Figure 1.2:

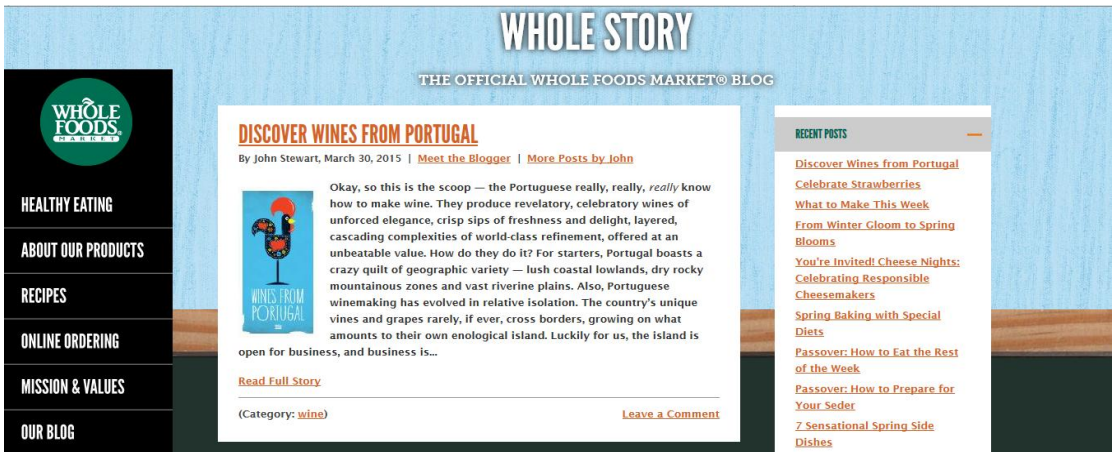


Figure 1.1: The corporate blog titled wholestory-1

(http://www.wholefoodsmarket.com/blog/whole-story?page=1)

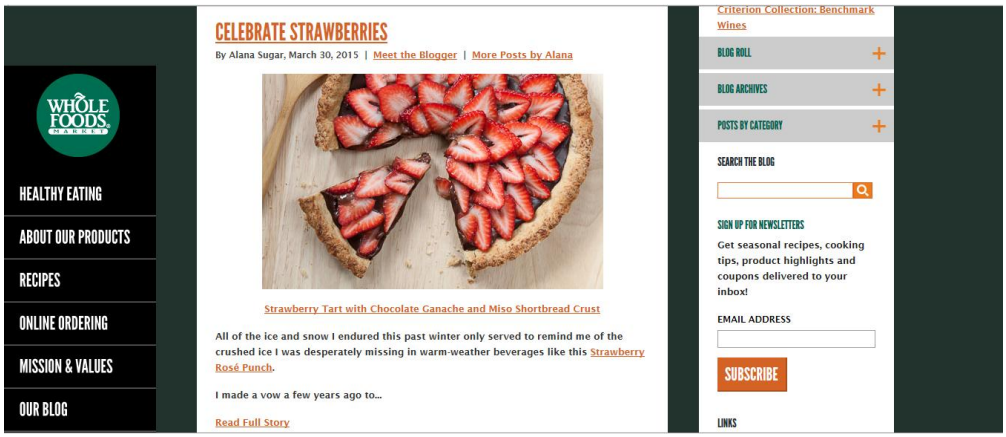


Figure 1.2: The corporate blog titled wholestory-2 (http://www.wholefoodsmarket.com/blog/whole-story?page=1)

In corporate blogs, the main elements include:

- An important area, in which the titles of some blog posts are listed chronologically, with the latest on top.
- Blog archives for older articles.
- A way for people to leave comments about the articles.
- A list of links to other related sites, sometimes called a 'blogroll'.

In the main content area, there are some articles (also sometimes called 'posts' or 'entries') that the author(s) writes. In a post or an entry, several component parts are included: a short, descriptive title, the body part which usually consists of both textual and non-textual content (visual or aural content), and the date and time it was posted. Some entries may be annotated with comments by the blog readers (Tirapat, Espiritu, & Stroulia, 2006). All these exhibit the complexity of blogs in structure, as well as its nature as a potential tool for a two-way communication between blog writers and their readers. The blog writer(s) writes an article, and the reader(s) often responds with a comment or idea in return. Whether responding to an article or not, blog readers can discover blog writers' characteristics in terms of their interests, their roles in the community, or their points by browsing their past blog entries (Nakajima, Tatemura, Hara, Tanaka, & Uemura, 2006). For both blog readers and writers, nobody needs to be experts when reading, commenting or publishing some content online; rather, the blogging software handles all aspects of formatting (Gillin & Schwartzman, 2010). Put it simply, the new technology behind blogging has created a form of new media for sharing information and knowledge between blog readers and writers. For corporate blogs, in particular, the blog

readers might be some current or potential customers, and writers are those on behalf of the entrepreneurs who might be hired freelance writers or even the staff from the company. The bloggers need to be professional and focus on themes their audience expects, which may include information directly related to the company itself, the writers of the blog, or the thoughts about the industry (Potts, 2007).

As a form of new media, the use of corporate blogs is consistently on the rise in some larger rather than smaller firms, such as *Marriott On The Move, Whole Foods, Ice Cream Journal, Delta Airlines,* etc. (McGregor, 2004). There are also some statistics showing the adoption of corporate blogs among the Fortune 500 from 2008 by the University of Massachusetts Dartmouth Center for Marketing Research (Barnes & Lescault, 2011). In 2013, 171 companies (34%) had corporate blogs, which is a great increase of the use of this tool since 2008, as can be seen in Figure 1.3. These include two of the top five corporations (Walmart and Exxon), leaving the other three (Chevron, Phillips 66 and Berkshire Hathaway) without a public-facing blog.

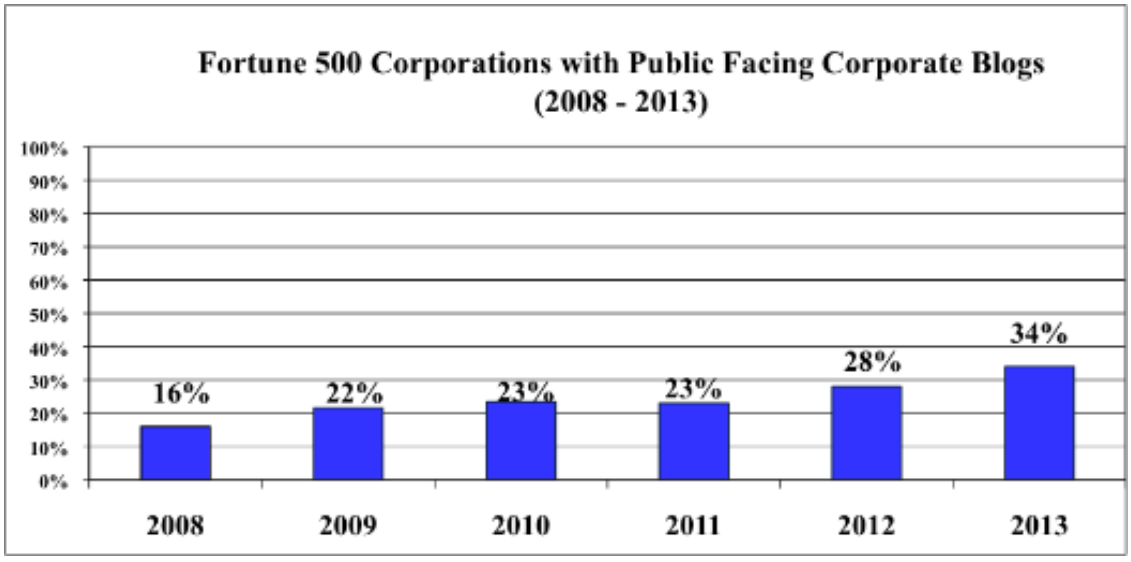


Figure 1.3: The use of corporate blogs since the 2008 study of the F500 (http://www.umassd.edu/cmr/socialmediaresearch/2013fortune500/)

The rise of the use of corporate blogs shows not only the needs of companies to reach their prospective customers on a person-to-person level, but also the situation that launching a corporate blog is able to benefit business companies. As indicated in Brecht, Cudreasova, and Zhou (2010), there is possibly a correlation between blog success characteristics, i.e., traffic

rank and connectivity, and the company success characteristics, i.e., sales rank. Besides, from less academic online sources, many internet articles are concerned about how to write a successful blog for business, in which some ideas are commonly agreed that blogs live by their content, particularly high quality content, and good corporate blog posts should speak to a target audience, tell stories to their audiences, and help their customers solve problems, etc. All these ideas have triggered me to think of another question, i.e., since the success of a blog could help its company to achieve success, then what content in a blog is able to make it successful.

Later on, I read a book named *The Language of Trust: Selling Ideas in a World of Skeptics* by Maslansky, West, DeMoss, and Saylor (2010), and it is after this that I finally associated the language of trust to top-ranked successful corporate blogs and formed an initial interest in exploring the language in corporate blogs. In this book, four principles of credible communication in the business world, i.e., 'be personal', 'be plainspoken', 'be positive' and 'be plausible', were proposed. Due to its basis on a wide range of research projects over the last decade, these principles are not only trustworthy, but also practical. What have particularly caught my attention are the personal and the plainspoken principle, which include several components, such as 'make it relevant', 'make it human', and 'make yourself real' (Maslansky et al., 2010). These principles and suggestions to either credible communication in the financial world or top ranked successful corporate blogs give rise to my curiosity to the language of corporate blogs, and the proposition of some relatively simple questions, as presented below: How personal is the language in corporate blogs? Speaking 'with' readers or 'to' readers? Whether or not there are any stories told in corporate blogs? Whether or not there are any instructions or information given to the audiences for solving some problems? With those relatively simple questions in my mind and based on my intuitive understanding of credible language in business communication and the correlation between the success of corporate blogs and the company success, I made an attempt to tentatively plan for my linguistic study on corporate blogs. The very starting point is an identification of the gaps in the literature, which helps formulate questions that need further investigation, as presented in the following section.

1.2 Gaps in the literature

As a relatively new phenomenon in the use of Internet and computer mediated communication, both of blogs and corporate blogs have not received much scholarly attention. In terms of linguistic-oriented studies, the blog has been considered as a hybrid genre that derives from other antecedent genres, and its language has much in common with spoken and written language (Aleknavičiūtė, 2009; Nilsson, 2003; Tavosanis, 2006) or with personal homepages and asynchronous discussion forums (Herring, Scheidt, Bonus, & Wright, 2004). In addition, distinctions between blogs and other genres have also been explored, such as marked variation between traditional academic genres and academic weblogs in the use of first- and second-person pronouns (Stuart, 2006) and formal and informal language use in blogs, email and school essays (Nowson, Oberlander, & Gill, 2005). An overview of the research on blogs is displayed by Myers (2010), who identified specific discourse characteristics and demonstrated the wide range of perspectives from which blogs can be analysed.

Previous research has also identified various types of blogs either based on their content types (Blood, 2002; Krishnamurthy, 2002; Kwasnik et al., 2005) or, less commonly, on their linguistic features (Grieve et al., 2011). In the study by Grieve et al. (2011), two main text types of blogs (personal blogs and thematic blogs) and a marginal third blog type (the expert blog) have been identified by using a multi-dimensional (MD) analysis (e.g., Biber, 1988, 1995, 2006; Biber & Kurjian, 2006), a comprehensive language-based analytical approach for quantitatively analysing registers and register variation.

As a specific type of blog, the corporate blog has attracted even less scholarly attention. Puschmann (2010a, 2010b) has pioneered research into this online language. In Puschmann (2010a), very broad contextual and extra-linguistic aspects of blogs and corporate blogs were examined, before the focus was directed to the linguistic aspects of corporate blogs. In this corpus-based linguistic analysis, it has been found that corporate blogs share lexical similarities (in the use of nouns and pronouns) both with the characteristics of business English and blogs in general, which “supports the broader claims made about the communicative qualities of corporate blogs” (Puschmann, 2010a, p. 110). In another study by

Puschmann (2010b), the linguistic properties of first- and second-person pronouns in corporate blogs have been explored, and several strategies for the use of such pronouns for self-reference and addressing audiences have been demonstrated.

Although these studies provide support for the view on blogs/corporate blogs as an emerging genre and important information about their linguistic features as well as functions, our understanding of the linguistic variation in corporate blogs remains limited in certain ways. Previous studies on corporate blogs focus on individual linguistic features; therefore, it is difficult to know how features interact and what the possible relations between forms (linguistic features) and functions are. Given these limitations, it would be important to adopt a more systematic perspective on linguistic variation in corporate blogs, from which a greater number of linguistic features as well as functional interpretations of language use can be explored. In order to accomplish these tasks, a theoretical framework is established at first, and then the research goals and research questions are specified.

1.3 Theoretical and methodological framework

This section provides a description of the theoretical and methodological framework for my research, in which a more systematic perspective is illuminated alongside the framework for this research, which centres on the register analysis, and combines both the MD approach to register variation, and the theory of SFL. These two components are compatible with each other due to the common grounds they shared. MD approach and the theory of SFL seem to be the practical end and the theoretical end of an entity of register analysis.

On the one hand, Biber's innovative work (e.g., 1988, 2006) in the field of quantitative study of register variation has had a major impact in linguistic studies. It not only includes quantitative, bottom-up analyses, but also introduces statistical techniques into linguistic interpretation. Methodologically, MD approach involves the use of multivariate statistical techniques (including factor analysis and cluster analysis) to study register variation in language corpus on the basis of the grouping/co-occurrence patterns of linguistic features and is directly applicable to understanding variation in corporate blogs. The identification of functionally

important linguistic features was largely dependent on many previous research studies. These background studies include previous comparisons of spoken and written texts (e.g., Chafe & Tannen, 1987), functional studies of particular linguistic features (e.g., Altenberg, 1984; Thompson, 1983), and descriptive grammars of English (Quirk, Crystal, & Education, 1985). This approach has played an important role in shifting researchers' attention from an individual analysis of particular linguistic features to viewing language variation as a continuum. A notable contribution of Biber's study is that it has strengthened the knowledge of vocabulary and grammar, as well as some communicative functions conditioning linguistic features.

On the other hand, it is sensible to adopt the theory of SFL for the current project, because in this way the communicative functions conditioning linguistic features could be more fully addressed. Besides, the combination of Biber's MD approach and SFL is based on the consideration that there are some common grounds shared, including functional view of language, co-occurrence patterns of linguistic features and the corpus-based study in linguistic research. With respect to the theoretical importance of linguistic co-occurrence, it has been emphasized by linguists such as Firth, Halliday, Brown and Fraser. Brown and Fraser (1979, pp. 38-39) observed that it can be "misleading to concentrate on specific, isolated linguistic markers without taking into account systematic variations which involve the co-occurrence of sets of markers". Halliday (1988, p. 162) defines a register as "a cluster of associated features having a greater-than-random...tendency to co-occur". This corresponds with the nature of Biber's MD approach, which identifies the major linguistic patterns of co-occurrence across the corpus quantitatively. Their difference is Biber provides a practical approach for the identification of text types that are distinguished linguistically, while other functional linguists tend to focus on the theoretical importance of linguistic co-occurrence.

When it comes to the functional view of language, it is also recognised in SFL. From an SFL perspective, language makes meaning in a social context, a context related to its social use and function. In terms of the meanings of language, Eggins (1994, pp. 78-79) claimed that language is "a social semiotic system organized as a set of choices and structures to make three kinds of meanings simultaneously". The three meanings include experiential,

interpersonal and textual meaning, and constitute the semantic stratum which acts as a crucial interface between context (the context of situation) and lexicogrammar. In the concept of register in SFL, the context of situation determines language use, and involves three parameters: field of discourse which specifies the topic of the linguistic exchange in the given situation, tenor of discourse which refers to the relationship between the participants in the situation, mode of discourse which describes the way in which the exchange is transmitted (Halliday, 1978; Halliday & Hasan, 1989; Halliday et al., 1964). With respect to the recognition on the functional nature of language, SFL provides a complex framework, encompassing not only language features, but some contextual factors labeling functional variation in register. Besides, the importance of corpus-based study in linguistic research has been emphasized by Halliday (1992), while the contribution of corpus-based study to the description of grammar is exemplified in Biber's corpus-based MD analysis.

The identification of the common grounds shared by SFL and Biber's work shows that it is feasible to follow a research procedure from an MD analysis to a corpus-based analysis of lexical and grammatical features in CB discourse, in which context of situation is realized. More theoretically, the two dimensions, i.e., instantiation and metafunction in SFL, also account for the relationship between text types explored in Biber's work and the context of situation in SFL. As shown in Figure 1.4, text types are systemic patterns of instantiation, located along the cline of instantiation at the intermediate region between the two poles of the cline - systemic potential and text instances, the relationship of which are like climate and weather (Halliday & Matthiessen, 2004). "By identifying a text type, we are moving along the cline of instantiation away from the text pole towards the system pole" (Halliday & Matthiessen, 2004, p. 27).

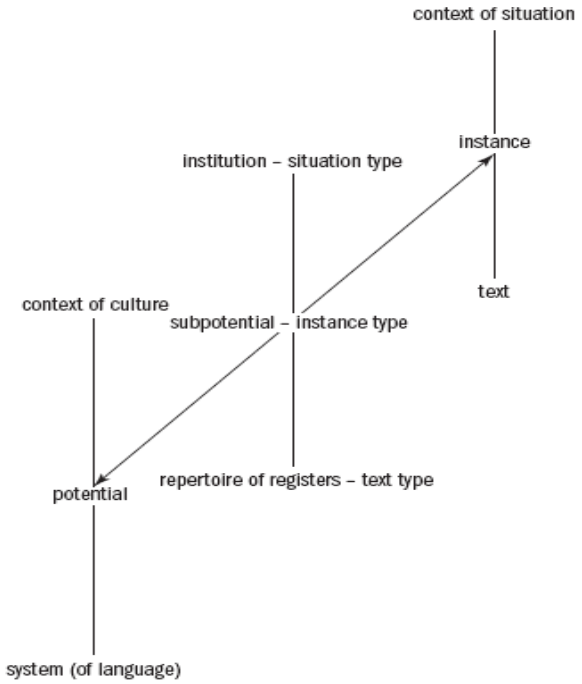


Figure 1.4: Location of text types (register) along the cline of instantiation (Halliday & Matthiessen, 2004, p. 28)

Besides, it should be claimed that the particular dimension immediately relevant to the more general study of text types in CB is that of metafunction. Presented in Table 1.1 are the five dimensions in language and their ordering principles.

Table 1.1: The dimensions (forms of order) in language and their ordering principles (Halliday & Matthiessen, 2004, p. 20)

Dimension		Principle	Orders
1	structure (syntagmatic order)	rank	clause - group or phrase - word-morpheme (lexico-grammar)
2	system (paradigmatic order)	delicacy	grammar - lexis (lexicogrammar)
3	stratification	realization	semantics – lexicogrammar - phonology - phonetics
4	instantiation	instantiation	potential - sub-potential or instance type - instance
5	metafunction	metafunction	ideational (logical-experiential) - interpersonal - textual

With this combined framework guiding the current research project, some assumptions could be made at first and then some research questions are raised. Firstly, the systemic concepts of Field, Tenor and Mode, and their related concepts on the dimension of instantiation can be applied in the analysis of different text types in CB. And the dimension of metafunction, as the meaningful realization of field, tenor and mode, also applies to all semiotic media. Secondly, it is assumed that the language of corporate blogs may represent an extremely complex discourse embedded in the highly varied institutional space of a particular industrial and business system. It hides a multitude of specific types of texts employed by various professional groups working in different contexts. The corporate blog spans a continuum of sources, ranging from service industries, to manufacturing, information technology, etc. (GICS[2] and ISIC[3] systems), as well as a continuum of different purposes from direct sales to branding. This indicates the diversity of corporate blogs, which could be divided into several broad categories.

1.4 Objectives and research questions

The theoretical framework constructed and discussed in the previous section paves the way for narrowing the research goals and raising research questions for this research. Overall, the current research project aims to:

1) reveal what are the principal dimensions of linguistic variation and the primary text types in corporate blogs;
2) reveal the relationship between CB text types and industries, i.e., whether or not the texts that are maximally similar in terms of their linguistic features are from a particular industry;

[2] The Global Industry Classification Standard (GICS) is an industry taxonomy developed by MSCI and Standard & Poor's (S&P) for use by the global financial community. The GICS structure consists of 10 sectors, 24 industry groups, 67 industries and 156 sub-industries [1] into which S&P has categorized all major public companies. The system is similar to ICB (Industry Classification Benchmark), a classification structure maintained by Dow Jones Indexes and FTSE Group. (http://www.msci.com/products/indexes/sector/gics/)

[3] The International Standard Industrial Classification (ISIC) of All Economic Activities is a United Nations system for classifying economic data. The United Nations Statistics Division describes it in the following terms: Wide use has been made of ISIC, both nationally and internationally, in classifying data according to kind of economic activity in the fields of production, employment, gross domestic product and other statistical areas. ISIC is a basic tool for studying economic phenomena, fostering international comparability of data, providing guidance for the development of national classifications and for promoting the development of sound national statistical systems. (http://unstats.un.org/unsd/cr/registry/regcst.asp?Cl=27)

3) and identify salient linguistic differences across text types in corporate blogs in the configuration of the three components of the context of situation - field, tenor, and mode of discourse.

In order to achieve these research goals, some research questions are formed, as the central issue being addressed in the current research. These research questions are located within the conceptual framework which combines both the MD approach and the theory of SFL. The proposition of them is based on an assumption that CB language represents a complex rather than a simple discourse, hiding a variety of specific types of texts employed by various groups working in different context. Among the three research questions, the answer to the first one relies on the use of Biber's (e.g., 1988) MD approach to register variation, which includes a factor analysis and cluster analysis, while a further and extended step to the typical MD approach can help answer the 2nd research question. Furthermore, by using the SFL approach, and from an SFL perspective, the 3rd research question is answered. The specific research questions are presented below:

Research question 1: What are the dimensions of functional linguistic variation and the primary text types in corporate blogs?

This question is tied to the 1st research goal, and answering it involves the employment of Biber's (e.g., 1988; 2006) MD analysis, which applies multivariate statistical techniques (including factor analysis and cluster analysis) to the investigation of register variation. On the one hand, the dimensions identified in MD analysis have both linguistic and functional interpretations. The linguistic content refers to a group of features that co-occur, with a markedly high frequency in texts (Biber, Conrad, Reppen, Byrd, & Helt, 2002). On the assumption that co-occurrence reflects shared functions, the co-occurrence patterns to assess the situational, social, and cognitive functions can be interpreted. By this means, this research question can be split into more detailed ones:

- What are the co-occurrence patterns of lexical and grammatical features in corporate blogs?
- What are the communicative functions of each group of co-occurring linguistic

features?

On the other hand, the identification of text types is based on an understanding that the text type is a variety of language that is defined exclusively by linguistic properties (Biber, 1989), which differs from register, a variety of language defined by situational (i.e. non-linguistic) characteristics (Grieve et al., 2011). That is, a text type is composed of texts that are maximally similar in terms of their linguistic features.

By addressing these issues, we are able to know how features interact, and the possible relations between forms (linguistic features) and functions in the corporate blog. In a more general sense, this study aims to enhance our understanding of how corporate blogs are constructed, as well as how corporate blogs are related to the universals of register variation. One perspective is exploring in what dimensions the variations of texts in any registers/genres basically vary; another perspective is finding out to what extent CB language is distinct from other discourse domain, reflecting the unique communicative and situational circumstances of CB.

Research question 2: How the text types in corporate blogs associate with the CB industry-categories?

This research question has a clear relation to the 2nd research goal. It can be considered as a follow-up of the first question. On the basis of the previous MD analysis, several text types are identified, any one of which is a variety of language that is composed of texts that are maximally similar in terms of their linguistic characteristics. Besides, in this research question, industry-categories are not well divided in accordance with linguistic features. Therefore, it is assumed that CB text types do not necessarily correspond to industry-categories. The exploration to the relationship between CB text types and industry-categories is basically a further investigation of linguistic form/discourse function relations in corporate blogs, in which industry-category is an added situational factor conditioning CB as a computer-mediated discourse. While industry-categories are defined based on situational characteristics, they tend to exert functional pressures on linguistic output by means of linking together with text types characterized by both functional and linguistic features in the other end. In this way, it

makes clear the linking between the linguistic features of CBs to their communicative functions, and possibly some unique style in communication and knowledge sharing between blog writers and readers. For example, when corporate blogs from wholesale and retail trade industry tend to tell stories, and give instructions rather than persuading their audiences to do something, they are more likely following some strategies for achieving the organizational purposes of corporate blogs. The results of this analysis will be presented in detail in the results section. Overall, this study tests and demonstrates a new application of MD analysis to validate linguistic realization of corporate blogs from different industries, an application which could be further extended to the validation of other genres and levels of discourse.

Research question 3: From an SFL perspective, what lexical and grammatical features distinguish different text types of corporate blogs, and how salient lexical and grammatical features serve to realize the meanings in the context of situation?

This question stems from the understanding of the shared common grounds between MD analysis and SFL in terms of functional view of language, and an agreement on the importance of co-occurrence patterns of linguistic features. They play a complementary role with each other theoretically and methodologically. It is assumed that the text type functions identified through an MD approach may encompass varied situational factors such as information about the participants, their relationships to one another, their purposes for communicating, and what they are communicating about (e.g., Baym, 1995; Hymes, 1974), which are largely overlapped by the parameters in the theory of context of situation. It is only when CB text types mapped onto Halliday's (2004) three metafunctions that the functional linguistic variation in corporate blogs can be understood more thoroughly. In this analysis, the language features at both lexical and grammatical levels are selected, allowing for an analysis of lexico-grammatical features in terms of ideational, interpersonal, and textual meaning. In particular, field of discourse specifies the topic of the linguistic exchange in the given situation, tenor of discourse refers to the relationship between the participants in the situation, and mode of discourse describes the way in which the exchange is transmitted (Halliday, 1978; Halliday & Hasan, 1989; Halliday et al., 1964). In order to achieve this, several major steps listed below should be followed: 1) Firstly, several linguistic features, including process type in transitivity, personal

pronouns, modals, lexical density, and grammatical complexity are selected from language metafunctions of mode, tenor and field within the SFL framework. 2) Secondly, the salient differences of the linguistic features mentioned above among different CB text types are identified, and then analysed in terms of its functional characteristics and contextual features. The language features selected from three metafunctions are at the lexicogrammar stratum that realizes the field, tenor, and mode in a particular context. I assume that the language features selected for the current research not only realize the strands of meaning, but also exhibit the grammatical and lexical meanings that should be two inherently connected parts of a single entity in its own right, rather than two separate domains (Sinclair, 1991). In this view, "a grammatical structure may be lexically restricted" (Francis, 1993, p. 142) and, conversely, lexical items are often grammatical in nature with some grammatical implications (Biber, Conrad, & Reppen, 1998; Conrad, 2000; Hunston & Francis, 2000). Taken together, in this system network, it is not only the three metafunctions but also the grammatical and lexical meanings that exist simultaneously in each sentence and correlate with each other. The meaning of a sentence is the product of both lexical and grammatical meanings, in which the contextual meaning is realized. It is not possible to separate the grammatical and lexical meanings that are two ways of picturing the same linguistic objective. For example, different process types, e.g., material process, mental process, etc., could be categorized into different grammatical categories in the semantic system networks, which is realized by the coordination of the lexis (meaning patterns) and the grammar (structure).

By addressing these issues, this analysis contributes to the existing field of knowledge in the following ways. Firstly, it develops the methodology used in corpus investigation of language variation, by means of extending MD analysis and identifying salient linguistic differences among different text types in an SFL framework. It also paves the way for further research into corporate blog and other forms of electronic communication. Secondly, it adds greatly to a description of corporate blog as a language variety in its own right.

1.5 Definition of key terms

1.5.1 Text

When it comes to the term *text*, it can be inevitably associated with another term *discourse.* The two terms have an intimate relationship with each other in literature. They can be used interchangeably, or sometimes differently, referring to written language and spoken language respectively. Some early linguists have preferred using one term rather than both of them, e.g., Halliday & Hasan (1976) and Quirk et al. (1985) prefer *text,* while Grimes (1975) and Sinclair and Coulthard (1975) basically use *discourse.* From early time to now, the two terms have been gradually coming closer to each other, as a result of the expansion of the scope of the notion text from a descriptive structural concept, towards a more process-focused concept that incorporates situational factors as well (Virtanen, 1990). For example, in Halliday & Hasan (1976, pp. 1-2), *text* is regarded as a semantic unit rather than a grammatical unit, referring to "any passage, spoken or written, of whatever length, that does form a unified whole". Basically, *text* is viewed as a *texture* (Halliday & Hasan, 1976), which distinguishes it from something that is not a text, as presented in many later definitions as well. Later on, the notion moves towards the process aspects of text and discourse production:

> "Discourse is a multidimensional process; 'a text', which is the product of that process, not only embodies the same kind of polyphonic structuring as is found in the grammar (for example in the structure of the clause, as message, exchange and representation), but also, since it is functioning at a higher level of the code, as the realization of semiotic orders 'above' the language, may contain in itself all the inconsistencies, contradictions and conflicts that can exist within and between such higher-order semiotic systems. Because it has this potential, a text is not a mere reflection of what lies beyond; it is an active partner in the reality-making and reality-changing processes" (Halliday, 1985, p. 318).

Overall, whether a text is spoken or written, it is a harmonious collection of meanings appropriate to its context. The term *text* in this project is used in reference to the main content

in a blog post which excludes other textual content, such as the title and the commentary part, and is distinguished from other content information, such as photos, audios, and comments etc. They serve as the individual units of data for the empirical analyses that the register studies in this research project are based on. The preferred units for the current research are include more than one complete text, and the compilations of complete texts could possibly prepare for the boundaries of different corpora. Overall, taking *texts* as the primary unit of linguistic analysis ideally accounts for a comprehensive register analysis.

1.5.2 Genre, Register and text type

In terms of the definition and the uses of *genre*, *register* and *text type,* in linguistics, there have been some controversial views on them, making them appear to be either a terminological maze (Moessner, 2001), or somewhat confusing and vague terms (Lee, 2001). Among the three terms, some attention has been given to the theoretical distinctions between *genre* and *register*. For example, Ventola (1984) and Martin (1985) refer to register and genre as different 'semiotic planes': genre is the 'content-plane' of register, and register is the 'expression - plane' of genre; register is in turn the 'content-plane' of language. The distinction between register and genre has also been applied in some linguistic studies; in these studies, genre tends to be associated more with the organization of culture and social purposes around language (Bhatia, 1993; Swales, 1990), whereas register is associated with lexico-grammatical features, showing how the use of particular words and grammatical features vary systematically in accordance with the situation of use (factors include interactivity, personal involvement, mode, production circumstances, and communicative purpose) (Halliday & Hasan, 1989). They can be regarded as two different perspectives looking at the same object.

The two terms have not been clearly distinguished in Biber (e.g., 1988, 2006); rather, they are simply used as a general cover term to refer to situationally-defined varieties described for their characteristic lexico-grammatical features. Whether in his previous or the latter studies, they are firmly rooted in the tradition of systemic functional grammar, in which *register* is "the clustering of semantic features according to situation type," and "can be defined as a configuration of semantic resources that the member of a culture typically associates with a

situation type" (Halliday, 1978, p. 111). Biber is also among those researchers (e.g., Lee, 2001) who distinguished *genre* from *text type.* Biber (1988) has mentioned that genre is based on external, non-linguistic criteria, while text type is based on the internal, linguistic characteristics of texts. As commented by Santini (2006), ever since the publication of Biber's work on linguistic variation across speech and writing in 1988, the term *text type* has entered corpus linguistics. His work on text typology by using an MD approach is considered to be "the most suggestive work so far in this area" (Lee, 2001, p. 40), and is by now "a classic of statistical corpus-based approach, and has influenced also European standards for large language resources" (Santini, 2006, p. 69). The textual dimensions proposed by Biber include more diversities, in comparison with traditional four rhetorical categories of narrative, description, exposition and argumentation by Faigley and Meyer (1983), or five text types (narration, description, exposition, argumentation and instruction) by Werlich (1976). Paltridge (1996) makes reference to Biber (1988) and considers the usage of 'text type' to be more concerned with discourse structures or patterns in texts, rather than lexico-grammatical or syntactic features. In this way, his 'text type' is similar to 'genre' proposed by such scholars as Hoey (1983), and Hatch (1992). Another researcher named Stubbs (1996) uses text type and genre interchangeably, without distinguishing them clearly.

In the current research, the corpus-based analyses rely implicitly or explicitly on the notion of text type and the related concepts register, since the text category is an important organising principle of the corpus. The register of corporate blogs is explored in terms of its linguistic features which relate the text to the aspects of the situation. This is a process of register specification, with the exploration of linguistic features as the point of departure, and then going on to make aware of the contextual situation guided by a combined framework encompassing both Biber's MD analysis and Halliday's SFL theory.

1.5.3 Narrativity and narratives

With respect to text types, the narrative is considered as a more fundamental and basic text type than descriptive, expository or argumentative text types (Biber & Finegan, 1989b; Virtanen, 1992). This is a text-type perspective on narrativity, which consists of "the formal and

contextual qualities distinguishing narrative from non-narrative, or marking the degree of 'narrativeness' in a discourse; the rhetorical principles underpinning the production or interpretation of narrative; the specific kinds of artifice inherent in the process of narrative representation" (Phelan & Rabinowitz, 2008). This is also a wider perspective to view narrative as a genre, which is a social activity fitting the context in which it occurs. Scale (2002) has suggested a far broader notion of narrative,

> I understand narratives to be constructed through many things, including acts of consumption, for example, which can be made symbolically to tell stories about tastes, relationships (whether real or desired) or social standing. (Scale, 2000, P. 37)

This is in contrast with those traditional ones, e.g., the narrative is a chronologically ordered representation of a series of events (e.g., Chatman, 1980; Genette, 1983) where events are specific time and place; the narrative refers to the devices and strategies governing the organization of a story into sequence (Forrester, 1996). If comparing narrative and narrativity, narrativity is a matter-of-degree, showing to what extent texts and speech are narrative. A wish for analytic clarity does not imply that narratives would exist as pure and distinct objects. It would be hopeless and misleading to assume that narratives are formally similar, always complete, and always neatly distinct from other kinds of discourse (Ochs & Capps, 2009). Both of narrative and narrativity are an umbrella term encompassing several different approaches to studying narrative communication (Koenig Kellas, Baxter, & Braithwaite, 2008). The earliest work is Aristotle's application of logical and ordered reasoning to the investigation of narratives in his Poetics, the purpose of which is to identify their different structures and components (Chatman, 1980). Another trend of work is the structural analysis of narrative with linguistics itself as a foundation. BarthesBarthes (1966) has identified three hierarchical levels of narrative linked by a progressive integration mode; Functions, Actions and Narratives. A function is a unit of content, each function being either distributive or integrative. The distributive class of functions is separated into two sub-classes of narrative units: the cardinal functions (core, articulation of the story) and the catalysis functions (to fill in the 'blanks' in the narrative space). Barthes also identified a set of two sub-classes in the integrative class: feature-based units and informants. Feature-based units are implicit and continuous, their role inside a story is to establish or amplify behaviours, feelings, atmospheres or philosophies;

informants help the identification and location of time and space. Feature-based units imply a descriptive activity (i.e., acknowledgement of behaviours or atmospheres) and informants usually bring knowledge and help to fix fiction into reality.

The standpoint of the current research is that narrativity is expressed by means of certain linguistic features. Biber (1988, p.109) distinguishes between narrative, or "active, event-oriented discourse" and non-narrative discourse, or "more static, descriptive or expository types of discourse". Multi-dimensional analysis of Somali, Korean and English all show a narrative dimension characterised by the past tense and temporal features that distinguish Fiction and traditional stories from other registers (Biber, 1993, p.341). This review paves the way for an exploration of the issue whether CB text types have an intermediate focus on narrativity or not.

1.5.4 Dimension

Dimension is an important element in Biber's MD analysis (usually including factor analysis and cluster analysis) (e.g., Biber, 1988, 1995, 2006; Biber & Kurjian, 2006), an empirical approach for the measurement of register variation by using multivariate statistical tools. In this approach, the linguistic co-occurrence is considered to be essential, and 'dimensions' are actually the co-occurrence patterns/factors that can be identified by using factor analysis, which reduces large number of features to a small set of factors (the term 'dimension' is exchangeable with 'factor' in this thesis). On the basis of the purely quantitative factor analysis, the linking of dimensions to functions is open to interpretation, and then a dimension/factor can be named, and further used as predictors in a cluster analysis. A grounding breaking work was done by Biber (1988), in which a broad range of spoken and written registers were analysed and several dimensions associated with communicative functions were identified: Dimension 1: Involved v. Informational; Dimension 2: Narrative v. Non-narrative; Dimension 3: Explicit v. Situation-dependent; Dimension 4: Persuasive; Dimension 5: Abstract v. Non-abstract. (cf. Chapter 2). No matter what dimensions are possibly identified to situate the registers in the current research, I will conservatively refer to the actual categories of variation analysed. In the course of the interpretation and the naming of dimensions in an MD analysis, it refers to a

large number of early studies or theories from different sources, such as Werlich[4] (1976) who belongs to another line of linguistic studies, i.e., German text linguistics which has not filtered into the English linguistic tradition. Thus, the sensible way to do is to follow the methods strictly but interpret the factors and dimensions comprehensively.

1.6 Outline of the thesis

This thesis includes seven chapters. In addition to the current introduction (Chapter 1), this thesis includes Review of the related literature (Chapter 2), Selection of linguistic features in the multi-dimensional analysis (Chapter 3), Corpus Design and Methodology (Chapter 4), Results of multi-dimensional analysis of corporate blogs (Chapter 5), Results of an SFL-based analysis of lexico-grammatical features of corporate blogs (Chapter 6), and the conclusions emerging from the analyses conducted (Chapter 7).

In Chapter 2 (Review of the related Literature), I present an overview of register theories, from the origins of register variation, to register in SFL. In the review of register in SFL, language description is located with respect to the dimensions of stratification, metafunction and instantiation. And then a stratified model of context of situation from SFL, which contextualizes linguistic choices with respect to the widely used systemic categories/parameters of field, mode and tenor, is presented. After this, an in-depth review of some language features related to or as a part of three metafunctions (field and the ideational function, tenor and the interpersonal metafunction, mode and the textual metafunction), i.e. the most common verbs and their process types, personal pronouns, modal verbs, lexical density and grammatical complexity, is given. Besides, I also review some previous studies on register variation by using multi-dimensional analysis, as well as comparing the kinds of dimensions identified in these studies. In this way, the present study of register variation is able to be connected with related lines of inquiry into the linguistic differences in a register, including blog register. Finally,

4 Werlich's (1976) five text types are very intuitive and reflect cognitive processes: perception in space (description), description in time (narration), comprehension of general concepts (exposition), creation of relations among concepts through the extraction of similarities, contrasts etc. (argumentation); planning of future behaviour (instruction).

I focus on the strengths as well as some weaknesses of both SFL and Biber's MD approach to register variation. A review from this perceptive is able to provide some supporting evidence, i.e., to what extent SFL and the MD approach share some common grounds and play a complementary role to each other, for the constitution of the theoretical framework for the present study. Taken together, all the works reviewed help to provide a framework for the present study.

In both Chapters 3 and Chapter 4, the methodological background is focused on. Chapter 3 presents a detailed review of the language features selected for the MD analysis, while Chapter 4 provides a detailed description of the methodological rationale and implementation of MD analysis, analysis of the industry, and SFL analysis as well as specific details of data collection and corpus design.

In Chapter 5, I present the results of an MD analysis of corporate blogs and then the results are evaluated in relation to some dimensions found in previous studies. In this study, six major factors/dimensions and six clusters/text types[5] are identified through factor analysis and cluster analysis respectively, and these are interpreted as dimensions of variation based on the shared functions of the clusters of co-occurring linguistic and discursive features. Also within this chapter, qualitative excerpts from representative texts are used to illustrate how these co-occurring features are indicative of a given function. This chapter also reports the distributions of different text types in CB across industries by using some mathematical measuring techniques. At first, the corporate blogs adopted in the corpus are categorized into 8 industries, and then the amount of CB texts contained in each text type, in each industry, and in each text type across 8 industries is calculated. On this basis, it is further revealed how each text type from each industry compares to the average level of the whole data.

In Chapter 6, some lexical and grammatical features associated with field, tenor and mode, i.e., the three register components in the setting of online corporate blogs are explored, and analysed at the discourse level in light of ideational, interpersonal, and textual metafunctions. It is also discussed how these findings echo the previous researchers' viewpoints, and how

[5] The two sets of terms are interchangeable: factor and dimension, cluster and text type.

field, tenor, and mode of CB discourse vary in a continuum. Finally, Chapter 7 summarizes the findings of this research as well as some methodological issues raised in the study. Besides, some suggestions and implications for further investigations are offered.

Chapter 2: Review of the related literature

2.1 Introduction

At the heart of the current study is an attempt to explore the variations in corporate blogs systematically from a perspective focusing on not only its linguistic features but also functions. This relies on an actual situation that the corporate blog is a form of communication that involves the control of lexical and grammatical constituents in social and situational contexts. Within the current research project which aims to provide a comprehensive linguistic description of variations among top-ranked corporate blogs, several approaches can be employed for answering the research questions raised in this project; they are Biber's MD analysis, and Halliday's register analysis based on field, tenor and mode. These approaches play different, as well as complementary roles to each other in this combined framework of register analysis. On the one hand, Biber's MD analysis can be considered as a very practical and empirical approach to the identification of dimensions or text types. On the other hand, Halliday (1988), together with some other linguists, such as Firth (1957), Ervin-Tripp(1972), and Hymes (1974), have emphasized the importance of some central elements in Biber's work theoretically, including the functional view of language, co-occurrence patterns of linguistic features and the corpus-based study in linguistic research. Besides, the relationship between text types and the context of situation is also claimed explicitly and systematically in the dimension of instantiations in Halliday and Matthiessen (2004). This provides a solid foundation for the consistency of two component parts of the analyses in this project, i.e., the investigation of CB text types by using the MD analysis and the exploration of field, tenor and mode-related language features across CB text types. Overall, the two components in the theoretical framework are based on the same perspective on language use, though representing different standpoints. In this chapter, the theoretical origins of register variation and register in SFL are both reviewed, and their strengths and weaknesses are also discussed.

In this chapter, motivated by filling the gaps found in the previous studies in blogs and

corporate blogs, many associated early works are reviewed and discussed, each of which is highlighted with the topics or research questions addressed, the approaches and methods employed, and the results obtained. In general, it starts from the review of the Hallidayan line of register theory, and then moves to the review of some previous studies on register variation by using multi-dimensional analysis, and some previous studies on blogs and corporate blogs. In particular, the review of the Hallidayan line of register theory provides the theoretical framework that is based on the dimensions of systemic functional theory. This is the foundation on which many implementational decisions are made and analyses are conducted. As crucial dimensions of the Hallidayan techniques for text analysis, the cline of instantiation, together with the hierarchy of stratification, and the spectrum of metafunction support the realization of patterns of meaning in texts (or discourse semantics) and clauses (lexicogrammar). This provides the background for the selection of language features for answering question 3 (From an SFL perspective, what lexical and grammatical features distinguish different text types of corporate blog, and how salient lexical and grammatical features serve to realize the meanings in the context of situation?). Closely related with question 1 (What are the dimensions of functional linguistic variation and the primary text types in corporate blogs?), some previous studies on register variation by using multi-dimensional analysis are reviewed. More specifically, in Section 2.2, there is an overview of the theoretical origins of register variation. Following this, Section 2.3 reviews register in SFL; sub-sections include Section 2.3.1 which is on stratification, metafunction and instantiation, Section 2.3.2 on register/context of situation which includes the discussion of three variables of register: field, tenor and mode; and Section 2.3.3 on lexicogrammar, in which language features related to three metafunctions are reviewed in Section 2.3.3.1, language features related to field and the ideational function are reviewed in Section 2.3.3.2, and language features related to mode and the textual metafunction are reviewed in Section 2.3.3.3. This is followed by a review of register variation and Biber's MD analysis in Section 2.3.4, and a discussion of the strengths and weaknesses of SFL and MD approach in Section 2.3.5. After this, a review of research on blogs and corporate blogs is presented in detail in Section 2.4, which provides a general background for the work investigated in this paper.

2.2 Theoretical origins of register variation

In linguistics, the most influential body of work on 'register' stems from what we refer to as "British contextualism" (Eggins & Martin, 1997, p. 237) represented by researchers from 'London school': the anthropologist Bronislaw Malinowski (1923), and his followers: Firth (1957), and his student, the British Australian linguist Halliday (1978), who founded systemic functional Linguistics. SFL is a sophisticated theory, which mainly consists of the assumptions that language is a system for meaning potential and that it performs a variety of functions. These assumptions basically stemmed from the ideas from Firth of London school, and Louis Hjelmslev of the Prague School (Halliday, 1983).

In Firth's work, many ideas are actually taken from the anthropologist Malinowski (e.g., 1923): a fundamental idea from Malinowski concerns the functions of language, i.e., language is functional, with certain functions performing in society, and thus implying a close relationship between language and society; another is the concept of 'context of situation', a description of the contextual factors influencing the text. This concept has been adopted and was even further developed, as stated by Halliday (1978, p. 109),

> Malinowski's notions were further developed and made explicit by Firth... who maintained that the context of situation was not to be interpreted in concrete terms as a sort of audiovisual record of the surrounding 'props' but was, rather, an abstract representation of the environment in terms of certain general categories having relevance to the text.

Apart from laying stress on the functions of language, and the context of situation, a 'system' perspective on language is also adopted in SFL. 'System' is a key notion in Firth's (e.g., 1957) work, which refers to a set of linguistic choices in a specific linguistic context. He has put emphasis on distinguishing the "paradigmatic" (system-based) description and the "syntagmatic" (structure-based) description (Halliday & Martin, 1981, p. 19), as well as calling for a framework of linguistic study that encompassed the polysystemic nature of language (e.g., Firth, 1957). In another line, the realizational view of language from Hjelmslev's (1961) work has also been adopted by SFL. As commented by Halliday (1978, p. 42), in the realizational view, language can be seen as "one system coded in another and recorded in another"; and in

the combinational view, “language is seen as larger units made up of smaller units”.

Rooted in anthropology and sociology, SFL basically adopted the functional view on language proposed by Malinowski, as well as the systemic and realizational views on language by Firth and Hjelmslev. However, it relied on a later development of Firth's work by Halliday and his colleagues that a more rigorous model is formed. Firth’s theoretical work which focuses on the areas of phonology and semantics was considered to be limited, while Halliday’s model has put more emphasis on the area of syntax, as well as contexts at the macroscopic level, which was assumed to be more appropriately engendering the study of register variation (e.g., Beaugrande, 1993). In such a trend, a more precise notion encompassing several elements including ‘contexts’ is developed inevitably. ‘Register’ is this precise notion, and it was Halliday, a pupil of Firth’s, and his colleagues eventually gave currency to this term.

In retrospect, the term ‘register’ was first used in Reid’s 1956 article for “text variety”, and further developed by Ure (1969, p. 107) as “situationally-differentiated language variety”. Some possible equivalents for it could even be found in earlier foundational linguistic works, such as “the universe of discourse” by Pike (1954) and “restricted language” in Firth’s work (e.g., 1937, 1968). Centring around language variety, linguists have been inclined to divide language variations into two distinctive dimensions: user-dependent, and use-dependent (Halliday et al., 1964). Examples of user-dependent variation include dialect variation, gender variation, etc. ‘Dialect’ is a type of user-dependent variation, which focuses on the study of language variation in “speech communities” (Milroy & Milroy, 2005, p. 267), and much work has been done on variation associated with social or demographic characteristics of speakers. In contrast, the use-dependent variation is more likely to be associated with “the language styles used by the same speaker on different occasions” (Milroy & Milroy, 2005, p. 275). Besides, use-related varieties are often associated with the term ‘register’, which is often more specifically referred to as “a variety according to use” (Halliday et al., 1964, p. 77), or a variety “in terms of language use” (Widdowson, 1983, p. 10). The term ‘register’ has been proposed inconsistently from linguist to linguist. In the early concept of ‘register’ by Halliday et al. (1964, p. 87), not only the position of use-dependent variation, but also a close relation between the

linguistic structures of a text and some elements of context could be reflected:

> "When we observe language activity in the various contexts in which it takes place, we find differences in the type of language selected as appropriate to different types of situation".

In comparison to this, more emphasis is put on the elements of context in Halliday's later definitions. As listed below, 'register' has also been defined as:

> "the set of meanings, the configuration of semantic patterns that are typically drawn upon under the specified conditions, along with the words and structures that are used in the realization of these meanings" (Halliday, 1978, p. 23).

Or,

> "a configuration of meanings that are typically associated with a particular situational configuration of field, mode, and tenor... the expressions, the lexico-grammatical and phonological features, that typically accompany and realize these meanings... [are] indices in the form of particular words" (Halliday & Hasan, 1989, p. 38).

Following Halliday's work, Gregory and Carroll (1978, p. 64) have commented that "register...is a useful abstraction linking variations of language to variations of social context...Register is an instance of language-in-action...Register is as well the realization of the semantic possibilities of language." In other words, register reflects the use of language which only has meaning in the situational or social context. Register-based variation is not determined by the language system itself because "register...is culturally determined since it is the culture of a society which determines the patterns of environments in which language can occur" (1978, p. 64). These patterns are acquired by the speaker through his or her own experience, which means that a certain predictability of register exists "which can...be described in terms of phonological, lexical and grammatical indexical markers (peculiar to a text) and common-core features (shared by texts)" (1978, p. 64). Besides, it was also pointed out by Gregory and Carroll (1978, p. 27) that there is a constant interaction between speakers and the situation in which they find themselves: "let us not forget that what we say is an indication of who we are as individuals, although even as unique persons our habits are neither fixed nor stable but mirror the constant variability of environment and attitude which makes up our lives". Similarly, Ellis and Ure (1969, p. 252) have also related text to context,

and defined register as “a linguistic category, a property relating a given text, in terms of its formal, phonological or graphological, or substantial; features to similar texts in comparable situations, and thereby to features in the situation of utterance or composition”. Both of the situations in Ellis and Ure (1969) and the patterns of environment in Gregory and Carroll (1978, p. 64) are further interpreted in Chilton (1978), who called them situation types. As its name shows, situations can be classified into several types: “a person in the social process makes a generalization about any particular situation, which enables him to see it as different, the same, or rather, in some other way related to other ‘situations’ he has experienced” (Chilton, 1978, p.114). He believes that a situation type could be “recognized as a result of a process of generalization facilitated by the existence of both linguistic and non-linguistic codes” (1978, p.115), which means it is not the situation alone but the interaction and cooperation between two codes that determine the register. In fact, he uses the term *register* to describe the ‘variation in language use’, which can be distinguished from accent and dialect variation. In this sense, Chilton is on the same line with Halliday as well as Gregory and Carroll, since all of them believe that register is closely linked to the “social meaning” (Chilton, 1978, p.123).

Moving from the theory to the practice, Chilton's concept of situation types and their relationship to register is also emphasized by Biber and Finegan (1994b) who actually conducted an empirical investigation to register variation by using the multi-dimensional analysis, a methodological approach that applies multivariate statistical techniques. In Ferguson (1994), it was emphasized theoretically that the underlying assumption in the study of register variation concerns “a communication situation that recurs regularly in a society (in terms of participants, setting, communicative functions, and so forth) will tend over time to develop identifying markers of language structure and language use, different from the language of other communication situations”. Biber and Conrad (2003, p. 175) also use the term *register* and describe it as “a cover term for any variety associated with a particular configuration of situational characteristics and purposes”. In Biber, Johansson, Leech, Conrad, and Finegan (1999, p. 15), some situational characteristics, such as “mode, interactiveness, domain, communicative purpose, and topic”, are related to register distinctions. Apart from the situational characteristics, several other components in register include the linguistic features,

and the functional relationships between the linguistic features and the situational or contextual characteristics (Biber & Conrad, 2009, p. 6). Among these components, the lexico-grammatical features appear in a target register due to their appropriateness for the situational context of that register. A register analysis consequently focuses on the pervasive linguistic features, which can occur in any variety but occur particularly frequently in the target register because of their connection to the situational context (Biber & Conrad, 2009). Although "any utterance is simultaneously determined by or exemplifies all four kinds of variation (dialect, register, genre and conversational variation)" in practice (Ferguson, 1994, p. 25), and register may be viewed on the same level as genre, the perspective adopted in a register analysis is register rather than genre.

2.3 Register in SFL

As indicated in the previous section, the development of the notion of register reflects a dynamic nature of register theory, as well as its essential position in SFL. On this basis, several register-related issues in SFL are reviewed in this section. What follows is a review of the location of lexicogrammar with respect to stratification, the dimension concerned with the global organization of semiotic systems into levels or strata; it subsequently moves to the functional diversification of clausal resources in terms of three kinds of metafunctions: the ideational, the interpersonal and the textual; finally it closes with a view on instantiation, the dimension whereby lexico-grammatical resources can be seen from two complementary perspectives, that of the overall potential available to the speakers/writers of a language.

2.3.1 Stratification, metafunction and instantiation

Before an overview of several crucial dimensions in SFL, a brief introduction to the following terms is presented at first: systemic, functional and social semiotic. Defined by Halliday and Matthiessen (2004), these terms are closely associated with the two basic theoretical postulates of SFL:

(i) The structure of language depends on its function of the communication medium in a social system;

(ii) Fundamental components of language structures are meanings. (e.g., Halliday, 1978; Halliday, 1994).

It is argued that SFL, on the one hand, is based on a systemic theory, which "gets its name from the fact that the grammar of a language is represented in the form of system networks, not as an inventory of structures" (Halliday & Matthiessen, 2004, p. 23). That is, language can be seen as consisting of sets of resources or options for making meaning, rather than sets of semiotics, such as computer languages. On the other hand, grammar is described as functional since the fundamental components of meaning in language, i.e., ideational, interpersonal and textual, are functional; and for any element in a language, its function in the total linguistic system can be indicated (Halliday & Matthiessen, 2004).

Besides, in a very early work by Halliday (1978), language is viewed from the social-semiotic perspective, i.e., semiotic systems are related to the meaning-making functions they serve, which particularly focus on the social functions that determine what language is like. It was also argued that the way to understand how language functions in relation to social structures is through the study of texts, the interpretations of which rely on the texts in the context. Accordingly, a text is not only an instance in its own right, but also an instance of social meaning in a particular context of situation. More specifically, "it is a product of its environment, a product of a continuous process of choices in meaning that we can represent as multiple paths or passes through the networks that constitute the linguistic system" (Halliday & Hasan, 1989, p. 11). In other words, texts can be considered as the "authentic productions of social interaction" (Eggins, 2004, p. 22), i.e., language produced in a particular context and used for particular social purposes. What follows is a brief review of several crucial dimensions in this multi-dimensional system: stratification, metafunction and instantiation.

2.3.1.1 Stratification

When it comes to the dimension of stratification, it means that language is a complex semiotic system including several levels, or "strata" (Halliday & Matthiessen, 2004, p. 24). The stratification theory of text in context have been proposed by many researchers from

Hallidayan school, such as Halliday (1994) and Martin and Rose (2003, 2008). Presented below is a stratified model of language from Butt, Fahey, Spinks, and Yallop (2000), which incorporates the extra-linguistic contexts (context of situation, context of culture) with language at linguistic levels, including content levels (semantics and lexicogrammar) and the expression level (phonology, graphology, etc.).

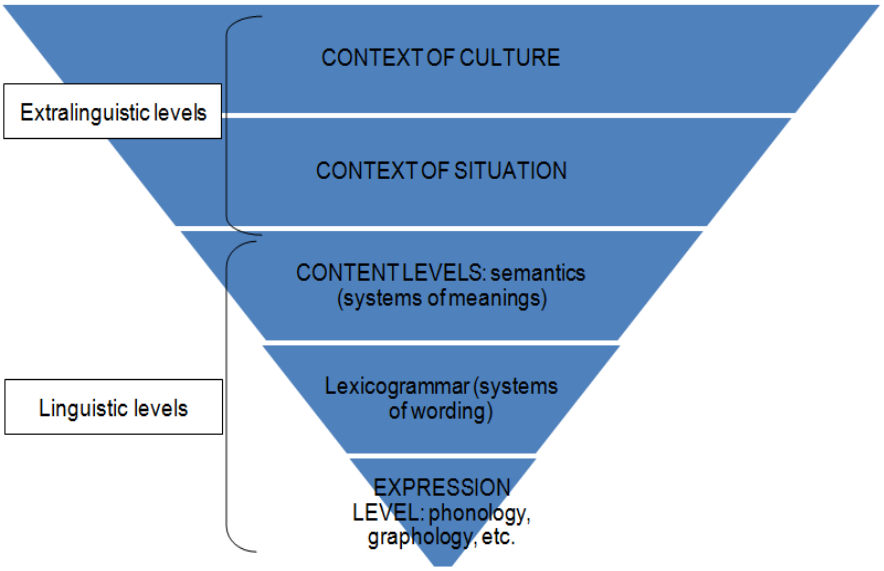


Figure 2.1: Levels of language - adapted from Butt et al. (2000, p. 7)

As shown in Figure 2.1, these levels are related by means of “realization”, which refers to the relationship among the strata, as well as “the process of linking one level of organization with another” (Halliday & Matthiessen, 2004, p. 26). It was also claimed in Martin and Rose (2008) that the layers of strata are related by means of realization, which allows patterns at different levels interpreted at another level. Similarly, in the stratification model by Halliday and Matthiessen (2004), language is considered to be a complex semiotic system consisting of various levels: context, lexicogrammar and a semantics at the content level, and phonetics and phonology at the expression level. The realization in these models is not one directional; rather, each of the layers could be realized by other patterns at any level. In comparison with these stratification models, Martin has provided an extended model, as can be seen in Figure 2.2.

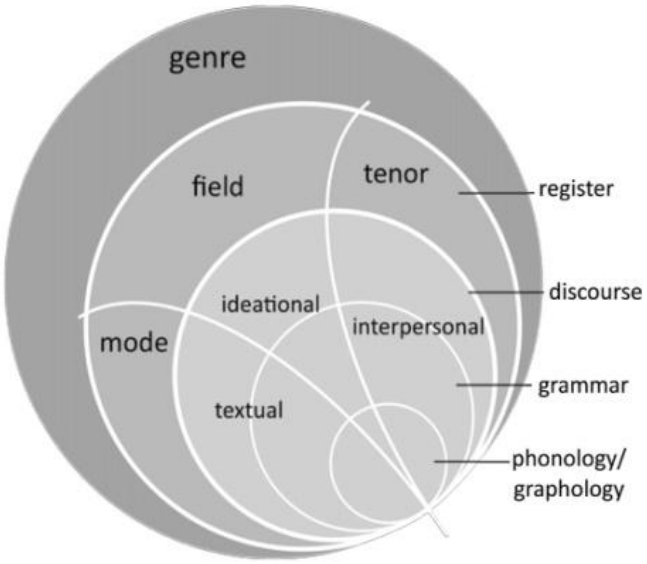


Figure 2.2: Language strata and metafunctions (based on Martin & Rose, 2007)

In this figure, the strata of language are divided into more stratification axis along its vertical dimension. More specifically, the functional model stratifies context into two layers: the context of culture and the context of situation. At the stratum of the context of culture, the recognizable patterns recur in texts which have similar social purposes and are identified as genres, a "staged, goal-oriented social processes" (Martin & Rose, 2008, p. 6). At the stratum of the context of situation, there are three aspects, in which language occurs in the social context: field, mode and tenor. This stratum will be focused on in this study, which is also called register, "a useful abstraction linking variations of language to variations of social context" (Eggins, 1994, p. 52)

In this model, the stratification of genre and register has allowed a multi-functional perspective on genre, cutting across register variables. Together, the three dimensions in the context of situation constitute the register of a text and are woven together to realize the genre. These contextual variables vary systematically in combinations and reflect the patterns of meanings. At the stratum of lexicogrammar, three metafunctions, which are basically social functions of language, are identified. They are the ideational metafunction that represents our experience to each other, the interpersonal metafunction representing our social relationships, and the textual metafunction which organizes the previous two into coherent and meaningful texts (Halliday & Matthiessen, 2004). The three meanings are in turn realized in the stratum of

lexicogrammar and phonology/graphology.

2.3.1.2 Metafunction

According to Halliday (e.g., Halliday, 1994; Halliday & Matthiessen, 2004), the three metafunctions of language, i.e., the textual, the interpersonal and the ideational meanings, are three separate strands of meaning, yet manipulated simultaneously and collectively. The ideational metafunction of language "provides a theory of human experience, and certain of the resources of the lexicogrammar of every language are dedicated to that function" (Halliday & Matthiessen, 2004, p. 29). In particular, it can be further divided into two component parts: experiential meanings and logical meanings. Experiential meanings in language are "the expression of the processes and other phenomena of the external world" (Halliday, 1978, p. 95), expressed through the system of transitivity or process type (Eggins, 2004), whereas logical meanings involve the mapping between these in language, such as the use of coordination, subordination, conjunction. Besides, the interpersonal metafunction is concerned with social role relations. The interpersonal function "serves to establish and maintain social relations, which include the communication roles we take on by asking or answering a question, and also for getting things done, by means of the interaction between one person and another" (Halliday, 1970b, p. 143). The textual metafunction is for not only construing experience but also enacting interpersonal relations, in the way as the resources for presenting information as text in context (e.g., Eggins, 1994; Halliday & Matthiessen, 2004).

Thompson (1996, p. 28) has described the three metafunctions in this way: "we use language to talk about our experience of the world, including the worlds in our own minds, to describe events and status and the entities involved in them; we also use language to interact with other people, to establish and maintain relations with them, to influence their behavior, to express our own viewpoints on things in the world, and to elicit or change theirs; finally, in using language, we organize our messages in ways which indicate how they fit in with the other messages around them and with the wider context in which we are talking or writing". Taken together, metafunctions would rather be seen as the functional components of a semiotic system, than merely the functions of the use of language. They could also be manifested as

three simultaneous strands or layers in the structure of the clause, which respectively construes different meanings in the physical and the mental world.

2.3.1.3 Instantiation

Instantiation refers to the relation between language as a system, i.e., an overall meaning potential, and text as a concrete instance (Halliday & Matthiessen, 2004). As stated by Halliday and Matthiessen (2004, p. 26), instantiation links system and text, and the relationship between system and text is a cline: text types are located along the cline of instantiation at the intermediate region between the two poles of the cline - systemic potential and text instances, the relationship of which are like “the weather and the climate”. By identifying a text type, “we are moving along the cline of instantiation away from the text pole towards the system pole” (Halliday & Matthiessen, 2004, p. 27) (cf. Figure 1.4). In the cline of instantiation (Halliday & Matthiessen, 2004, p. 28), system and text are located at the two poles of the cline - that of the overall potential and that of a particular instance. Between these two poles there are some intermediate patterns, which can be viewed from different perspectives. Viewed from the system pole, these intermediate patterns are “subsystems”, and viewed from the instance pole, they are “instance types” (Halliday & Matthiessen, 2004, pp. 27-28). When identifying a text type, we move along “the cline of instantiation away from the text pole towards the system pole” (Halliday & Matthiessen, 2004, p. 27). In practice, however, texts may vary systematically in accordance with the contextual factors.

Martin (2010) has also provided an elaboration to instantiation hierarchy, in which genre/register and text type are added as further steps in the degrees of generality (Figure 2.3 below). For him, instantiation is a hierarchy of generalization moving from all the resources available in the system to some generality - recurrent patterns, including register/genre, text type, and even the text. As a recurrent configuration of field, mode and tenor patterns, genre sits on top the realization hierarchy; however, it is between system and the text, since each genre constitutes a subpotential of the meaning potential of the system as a whole.

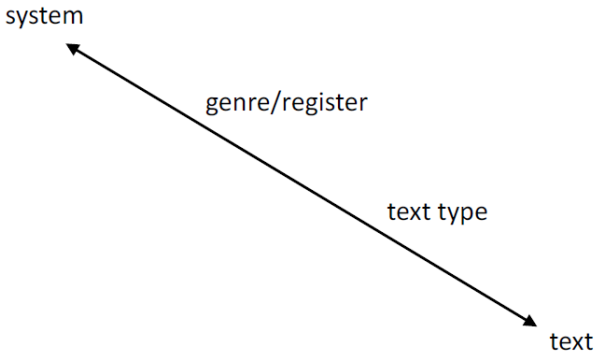


Figure 2.3: The cline of instantiation, including genre and registers as sub-potentials (based on Martin, 2010, p. 18)

This figure is very concise; however, it can be assumed that a complex system is behind it, in which many interrelated semiotic dimensions can be involved, such as instantiation, metafunction, etc. For example, when instantiation is mapped onto the strata, all strata along the realization hierarchy are able to instantiate. In contrast, without it, merely moving from genre to graphology/phonology cannot arrive at the instance of language use. Presented below is a more detailed figure with instantiation mapped onto strata, metafunction and rank (Figure 2.4).

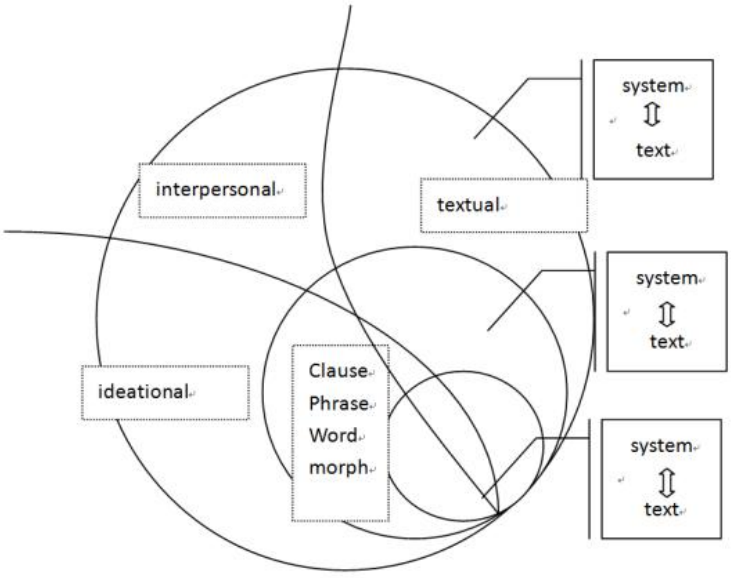


Figure 2.4: Instantiation mapped onto strata, metafunction and rank (based on Martin, 2010, p. 22)

In Figure 2.4, the system positions in the highest degree of generality, in which all types of the meaning potential are located. It is also shown that the relationship between system and instance is a dynamic one: the system evolves and expands through its constant instantiation

through an ongoing process, which can be seen along three time frames:

1. the evolution of language in the human species, or within a linguistic community;
2. language development of an individual throughout their lifetime, and
3. the meaning-making process in the unfolding of texts (Halliday & Matthiessen, 1999; Martin, 1999a)

All these different kinds of meaning-making processes make the system open-ended and dynamic. When it comes to lexicogrammar in the whole system, it can be viewed from two perspectives: the overall lexico-grammatical potential in a language, and the lexico-grammatical resources in situated texts. There exists some difficulty establishing a linking between the description of patterns and individual instances, since the description of patterns in texts may result in failure to generalize across instances, and the focus on individual instances may lead to the inability to relate specific configurations of resources to patterns across registers and genres. In spite of this, it can be feasible to relate instantiated texts to different patterns.

To sum up, instantiation plays an important part in language descriptions; it is usually located at intermediate points between the system and the instance in the cline, e.g. patterns in specific registers/genres that can be later used in comparative and/or contrastive studies (e.g., Matthiessen, Teruya, & Canzhong, 2008), and accordingly a balance is needed between the system and the instance. Besides, the importance of instantiation lies in the fact that it allows for the patterns located in texts as a source of “data” which relates to the generalizations of the whole system (Caffarel, Martin, & Matthiessen, 2004, p. 20).

2.3.2 Register: Context of situation

In this section, a stratified model of context of situation from SFL, which contextualizes linguistic choices with respect to the widely used systemic parameters of field, mode and tenor, is reviewed. These three contextual parameters are associated with their respective functions and purposes: meanings in tenor are realized through interpersonal metafunctional choices; field meanings are realized through ideational metafunctional choices, and, in the same manner, meanings in mode are realized through textual metafunctional choices (e.g., Halliday,

1994; Halliday & Matthiessen, 2004). In comparison with this, some alternative SFL models of context were developed by some other researchers as well, including Gregory's (1967) model consisting of field, mode, personal tenor, and functional tenor, Ure & Ellis's (1977) model consisting of field, mode, formality and role, and Fawcett's (1980) model consisting of subject matter, channel, relationship purpose, and pragmatic purpose. Although the category of register in these models are constant in general, the tension between these frameworks have been widespread (Young, 1985). What has been adopted in the current research is the classical stratification of register by Halliday and his colleagues (e.g., Halliday, 1994; Halliday & Matthiessen, 2004). This is based on the consideration that the current research is more concerned with the "practical utility of the concept of functional tenor" as well, and thus, Halliday's model which has more comprehensive correlation of metafunctions, including functional tenor, is adopted (Martin, 1999b, p. 27).

2.3.2.1 Field

According to Halliday and Hasan (1989, p. 12), the field of discourse refers to "what is happening, to the nature of the social action that is taking place: what is it that the participants are engaged in, in which the language figures function as an essential component". Similarly, Eggins (1994) has also considered field as an essential variable that is associated with the activity that we are engaged in. Taken together, the register parameter of field centres on the nature of social action, which involves some use of language.

Steiner (2004) summarized that this parameter includes several sub-dimensions, including experiential domain, goal orientation and social activity. In terms of the category of experiential domain, Halliday and Hasan (1989) call it the nature of the social activity, which makes the borderline between experiential domain and social action even more blurred. Their obscure borderline also lies in the fact that both of them could be concerned with rather abstract activities, such as production, exchange, communication, reproduction and consumption. Similarly, Martin (1992) proposes a classification of field by means of modeling field as sequences of activities, which also involve some entities, such as people, or things. In this field network, several aspects of fields are involved and distinguished, which are on the scale with

respect to “degree of specialization and technicality, alongside the scaling of oral and written modes” (Martin, 1992, p. 545). It is then commented by Martin (1992, p. 545) that “this kind of ranking should not be read into the networks at the level of field”, because fields are usually “valued in different ways and so ranked by all cultures” (Martin, 1992, p. 546). I also consider that this taxonomy is too abstract to be applied to an empirical study (see Figure 2.5).

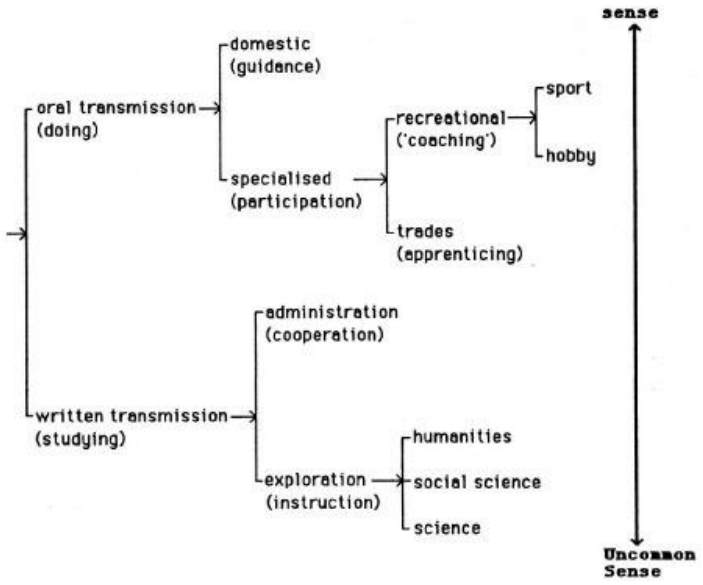


Figure 2.5: Martin's (1992, p. 544) classification of fields

Apart from this, the category of goal orientation is described in various ways by different researchers. Hasan (1999, p. 234) considers goal as “an inherent aspect of human social action and thus as an important component of a text’s relevant context”. She points out that this view has been accepted in SFL, whether in Martin’s semiotic model or the Hallidayan model. She also proposes several crucial considerations: social agents’ awareness of the goal/motive of their action is varied; goals/motives do not necessarily describe conscious mental states; goal of a social activity is inevitably related to the outcome. It is also in this line that Butt (2004) offers different parameters for the description of goals. He characterizes goal orientation as “an attempt to elucidate the (outward) indices of the inner controls and organization of meaning in a context” (Butt, 2004, p. 34).

In general, there are some commonly shared drawbacks of these theoretical works: it is not clear how these abstract parameters can be realized at linguistic levels, particularly at a lexico-grammatical level, although it should be admitted that theory is able to provide a

systematic view of a phenomena by specifying the relations among variables and propositions with some purposes. In spite of this, an insight into the combination of theories and empirical evidence may help overcome this problem. In order to explore the field of the text, the selection of a particular language feature, such as the most common verbs and their process types, rather than the inclusion of a wide range of language features related with the field, e.g., lexical fields, terminology, lexical chains, transitivity, headings (Steiner, 2004) is more feasible. What follows is a brief review of tenor of discourse that the researchers are more concerned about.

2.3.2.2 Tenor

The tenor of discourse is originally defined as "the cluster of socially meaningful participant relationships, both permanent attributes… and relationships that are specific to the situation, including the speech roles" (Halliday & Hasan, 1989, p. 143). In more detail, the parameter of tenor refers to

> "who is taking part, to the nature of the participants, their status and roles: what kinds of role relationship obtain among the participants, including permanent and temporary relationships of one kind or another, both the types of speech role that they are talking on in the dialogue and whole cluster of socially significant relationship in which they are involved." (Halliday & Hasan, 1989, p. 12).

Centring on the relationship between the speaker and hearer, some sub-dimensions in tenor have been proposed by different researchers. Tenor has been modelled as including three distinct features: agentive role, social hierarchy (status, power) and social distance (e.g., Hasan, 1979). Similarly, Poynton (1989) has also divided the tenor of discourse into three sub-dimensions: power, familiarity, and affective involvement. In Butt (2004), tenor has been modeled as parallel systems in a network, including two dimensions; they are agentive role, and social hierarchy.

There are some overlaps of these sub-dimensions. The agentive role in Hasan (1979) is similar to that in Halliday and Hasan (1989). Halliday and Hasan (1989, p. 56) have elaborated

that “the nature of the activity predicates the set of roles relevant to the unfolding of the activity”. In Hasan (1999, p. 247), there is an increasing awareness of the complexity of interactants’ roles in communication: “it is difficult to identify one single relation of the agentive kind which would apply constantly to the entire dialogue”. In view of this, the complexity of the mutual communication potentially requires a separate analysis, rather than the analysis of agentive roles as a whole, which is actually pointless. Butt’s (2004) model is an example, in which tenor is modelled as parallel systems, giving rise to more numerous contrasts between co-selections of field, tenor, and mode. The value of this approach is that its complexity may let the analyst map “the character of a particular instance (or set of instances) as a particular configuration of choices across all these ‘domains of contrast’ and compare it against any other instance, broader dataset/corpus, or idealized variety” (Lukin, Moore, Herke, Wegener, & Wu, 2008).

In terms of social distance and social hierarchy, they are more closely intertwined with each other. Social hierarchy is sometimes referred to as` social role relationship; it is concerned with the degree of control or power one interactant has over the other (Halliday & Hasan, 1989). The sub-dimension of familiarity in Poynton (1989) is concerned with the social distance which basically reflects the degree of familiarity between interactants (Hasan, 1979). Butt (2004, p. 16) characterizes the category of social distance as “classifying the extent of the relationship between the participants in terms of density as well as formality of context”. Formality, no matter a greater or a lesser one, reflects the relation between a speaker and the addressee(s) in a given situation. Joos (1967, p. 45) has even distinguished five stages of formality: “frozen, formal, consultative, casual, and intimate”. Some other researchers even contained some more concrete language features as observable indicators for this category, such as tone, accent, ellipsis, vocation and terminology from Martin (1992), tagging, forms of address, modality, accents, and dialects from Steiner (2004).

Besides, power refers to the relative positions of the participants in a social hierarchy, with equal and unequal status (Poynton, 1989). Similarly, the degree of power or control one interactant has over the other is named “social hierarchy”, or social role relationship (Halliday

& Hasan, 1989, p. 57). Social hierarchy gives us insights into whether the writer and the reader hold equal social roles or being in a hierarchical relationship. Previous studies show that particular pronouns signal the relation of the person to the listener (Brown & Levinson, 1987). The use and choice of pronouns provide important insights into social relationships because pronouns have linguistic functions that contribute to construction of identity and a social reality (Mühlhäusler & Harré, 1990). In SFL, Halliday (1978) claimed that the language features most closely associated with the interpersonal function are mood, polarity, modality, and personal pronouns. Modality could be placed as a part of the interpersonal function in that it expresses "a role relationship between the speaker and hearer" (Halliday, 1970a, p. 325). It is often referred to as "a form of participation by the speaker in the speech event", and it is often through modality that "the speaker associates with the thesis an indication of its status and validity in his own judgment", thus "intruding" and taking up a position (Halliday, 1970a, p. 335).

2.3.2.3 Mode

Mode has been defined as

> "what part the language is playing, what it is that the participants are expecting the language to do from them in that situation: the symbolic organization of the text, the status that it has, and its function in the context, including the channel (is it spoken or written or some combination of the two) and also the rhetorical mode, what is being achieved by the text in terms of such categories as persuasive, expository, didactic, and the like" (Halliday & Hasan, 1989, p. 12).

It often includes the role language plays in the realization of social action (Martin, 1992), as well as 'how the language is functioning in the interaction' (Thompson, 2004, p. 40). Lavid (1995) sums up that the mode of discourse includes three simultaneous parameters: language role, the medium, and the channel of discourse.

In terms of language role, Hasan (1999, p. 281) argues that this sub-dimension is more concerned with the field of discourse. Her main point is that "the so-called rhetorical modes such as explaining, defining, generalizing, reporting, recounting, narrating, chronicling etc. are

best viewed as constitutive verbal actions and thus as specifying the nature of the social activity which is accordingly analysed under field of discourse". In the present study, it is accepted that these rhetorical modes can be treated as goals of verbal action from the sub-dimension of field. Besides, Halliday and Hasan (1989) have pointed out that the sub-dimension 'channel' covers the heading 'medium' due to the differences of transmission and production in language, rather than the spoken-written distinction. However, I suppose this correlation is narrowly focused and may not necessarily apply to some more recently emerging language phenomena. For example, in a chatroom, some special cases are expected to emerge: people may talk like speeches, but they tend to communicate in a written form by typing on a computer keyboard. This has inevitably involved both the issue of medium, and spoken/written language distinctions. Therefore, I also adopt some claims by Martin (1997) and Eggins (2004). In this line, 'mode' is viewed as connecting to the semiotic distance that is affected by the communication channels through which we act socially (Martin, 1997, p. 12). Mode is concerned with symbolic reality so it is oriented to both interpersonal and experiential meanings: it mediates the role played by language along two dimensions - the interpersonal dimension (tenor), and the experiential dimension (field). Eggins (2004, p. 91) develops these concepts in two simultaneous continua, which describe two types of distance in the relation between language and situation:1) "interpersonal and spatial distance between the interactant (associated with feedback or its absence), and 2) experiential distance between language and the social process occurring (language as action or as reflection constructing experience, reflecting on action)". According to Eggins (2004), the continua of these distances characterize the basic contrast between spoken and written language. There are many language features characterizing spoken and written languages, such as a turn-taking organization, non-standard grammar, grammatical complexity, lexical density, etc. What have been selected for the current research are the features of lexical density and grammatical complexity, which are highly sensitive to medium (e.g., Neumann, 2010), as well as positioning between the extremes of the spoken and the written language.

2.3.2.4 Summary

In the brief overview of the three parameters in register presented above, the issues of the boundaries for models of context within SFL, and the particular sub-systems within each of these have been introduced. In relation to the work of the stratification of register, Martin (1999b) has raised some points concerning the solidarity and combinations of the field, mode and tenor. He confirms the solidarity of the filed, tenor, and mode proposed by Halliday. Without this solidarity, the elegance of SFL which explains how language is used seems to be seriously compromised (Martin, 1991). He also considers the association of register variables and metafunctions to be a good start for research. However, the potential danger is field, tenor and mode options may over-generate the meaning potential taken up by a given culture, if they are proposed as simultaneous social semiotic systems. Besides, Hasan (1999, p. 232) has commented the exploration of sub-systems in three parameters this way: “since language in use realizes some given CC (contextual configuration), any variation in CC will naturally activate some variation in this language; it is this kind of variation that in SFL we refer to as register variation.” These points suggest that the potential contrasts between co-selections of field, tenor, and mode could be great, thus the characteristics of a particular text or set of texts as a particular configuration of choices can be varied. However, there are some difficulties to move from theory to practice; this has been echoed in some critical points by some researchers. Hasan (1999) has raised a problem concerning the inclusion of situational features in the discussion of a text’s context, which reflects an uncertainty that exists in context modeling. She also claimed that “our description of context runs the risk of being as unmanageable as ’transcribing infinity’ as feared by Cook (1990)”. This may leave aside some questions as presented below:

> “on what basis can we justify the inclusion of all such details? More basic still: would we really know what the expression ‘all such details’ refers to? How do we conclude that enough has been said about some specific context? On the other hand, if to make context an effective tool for analysis it must be ‘contained’, then we need to be clear what aspects of the interactants’ material and social conditions of existence are integral to the concept and why?”(Hasan, 1999, p. 232).

The proposition of these questions reveals not only the wide varieties of language features, but also the issue concerning how concrete aspects of a situation can be derived. In practice, it is impossible to give a comprehensive account to any aspects in the language, including all systems, strata and ranks (Lukin et al., 2008). In spite of this, these 'contextualization systems have the distinction that instead of taxonomizing realized meanings, they actually systemize the realization-activating contextual features and attempt to relate context to wording via meaning, which acts as the interface between the two' (Hasan, 2009, pp. 181-182). Many plausible different co-selections are available, both within and across metafunctions. Martin (1992) has suggested that a hypothesis about the association of register variables and metafunctions is the better place to start a research, and moving from the field, mode and tenor plane to a more abstract level of genre may strengthen the hypothesis. He also mentioned people's interests in undertaking quantitative studies of register, which might be used as evidence for solidating register and metafunction, such as the methodology pioneered in Horvath (1985) and Biber (1988).

Therefore, I consider that the following two issues should be adopted in the framework guiding the current research: one is that particular attention should be given to the aspect which directly reflects the variation in the configuration of field, tenor, and mode, rather than all aspects in the discourse reservoir; the other is the issue of method which means register cannot be described and analysed simply on the basis of single instances. The selection of language features from three metafunctions is for an extended study of MD analysis of CB, i.e., a further step for answering research question 3 in this project. The criteria for the selection of language features are based on some previous works on register variation (e.g., Biber, 1995; Martin, 1992). 'The most common verbs and their process types' is a suitable language feature from the field of discourse which deals with the topic of the context where language is used. 'Pronouns' and 'modal verbs' are associated with the social role relationship in the tenor of discourse and 'lexical density and grammatical complexity' are from the mode of discourse which concerns the complexity of language, and the spoken/written distinction. Other language features associated with three metafunctions, such as tone, accent, ellipsis, vocation and terminology from Martin (1992), tagging, forms of address, accents, and dialects from

Steiner (2004), are not adopted because they are much less related with my research which is triggered by the principles of credible communication from Maslansky et al. (2010) (cf. Section 6.2). Presented in the following section is a review of the language features selected for the current research from the stratum of lexicogrammar.

2.3.3 Lexicogrammar

This section focuses on another stratum of language, i.e., lexicogrammar, which is the core stratum in Halliday's model of language referring to "words and the way they are arranged" (Butt et al., 2000, p. 6). It mediates between a lower level of "sounding" (graphology / phonology) and a higher level of "meaning" (semantics / discourse). The standpoint of this research is both lexis, and grammar since lexicogrammar is not a separate stratum from semantics. Then, the consequence is that the language, and in particular the lexicogrammar, is structured by the expressive and communicative functions it has evolved to convey. In the following sections, several linguistic features, as the component of three metafunctions, i.e., ideational, interpersonal and textual, will be introduced from a discourse analytic point of view. The rationale for this emphasizes the functional, semantic phenomena, as well as their forms that traditional grammar mainly focuses on.

2.3.3.1 Features related to field and the ideational function: transitivity

In this section, the process type of transitivity, which is the primary system within SFL for expressing the experiential meanings (ideational metafunction includes both the experiential and logical metafunctions) will be focused on. The main function of the grammatical system of transitivity, argues Halliday and Matthiessen (2004, p. 170), is to construe "the world of experience into a manageable set of process types". Process type is an element of transitivity, which is realized by the verbal group and also includes participant and circumstance. As commented by Thompson (1996, p. 77), the actual process, which refers to the event or state that is described, is "the central component of the message from the experiential perspective". The participants are the people, or objects associated with the process. And the circumstances provide contextual information for the process and are typically realized by adverbial and

prepositional groups.

The process types in transitivity have been divided into six categories: material, mental, relational, behavioural, verbal and existential processes (e.g., Halliday, 1994; Halliday & Matthiessen, 2004), as shown in Figure 2.6.

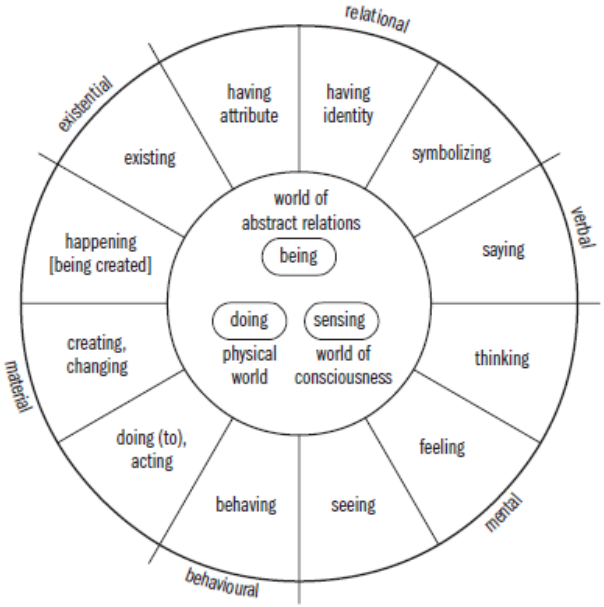


Figure 2.6: The grammar of experience: types of process in English (Halliday & Matthiessen, 2004, p.172)

Among the six process types identified, the material, mental and relational processes are the principal types, while the behavioural, verbal and existential processes are subsidiary types, which are located at the boundaries of the three major ones and share characteristics with them. As stated by Halliday and Matthiessen (2004, p. 248), those principal types are “the cornerstones of the grammar in its guise as a theory of experience; they present three distinct kinds of structural configuration, and they account for the majority of all clauses in a text”. To illustrate this, some examples are provided in Table 2.1.

Table 2.1: Examples of different process types (Halliday & Matthiessen, 2004, p. 171)

Process type	Example (Process + participants underlined; Process in bold; circumstances in italics)
material	*During the European scramble for Africa,* Nigeria **fell** *to the British.*
	and the British **ruled** it *until 1960*
behavioural	people are **laughing**.
mental	The Ibos **did not approve of** kings.
verbal	so **we say** → that every fourth African is a Nigerian
	Can you **tell** us *about the political and cultural make-up of Nigeria?*
relational	that every fourth African **is** a Nigerian.
existential	so *today* there**'s** Christianity *in the south*

To better understand process types, a brief overview of each of them is needed. Material processes are the ones most closely associated with doing, and happening in the concrete world; they are typically physical actions. The participant of a material clause is an Actor - the one who is performing the action, while what is affected by the action is named a Goal, which is usually realized by adverbial groups, prepositional phrases, and even by nominal groups.

Mental processes are concerned with our own consciousness and experiences of sensing, such as processes of affection (e.g. like, hate), cognition (e.g. know, understand) and perception (e.g. see, hear, feel), etc. A mental clause construes “a quantum of change in the flow of events taking place in our own consciousness. This process of sensing may be construed either as flowing from a person’s consciousness or as impinging on it; but it is not construed as a material act” (Halliday & Matthiessen, 2004, p. 177). The human participant involved in mental processes is often referred to as the Senser and what is sensed is called the Phenomenon.

Relational processes, or processes of “being”, relate a participant either to an identity (e.g. He is my brother) or to an attribute (e.g. He is good-looking), which means relational clauses may serve to characterize and to identify as well. There are two main types in this process: relational “attributive” and relational “identifying” processes (Halliday & Matthiessen, 2004, p. 216). In relational attributive clauses, the main participant is the Carrier and the characteristic assigned to the Carrier is called the Attribute. In relational identifying clauses, the main participant is the Identified and the identity assigned to it is the Identifier.

Behavioural processes have been considered as the least distinct of all the six process types because they have no clearly defined characteristics of their own; rather, they share characteristics of both material and mental processes, with typically "physiological and psychological" activities, such as breathing, coughing, smiling, dreaming and staring ((Halliday, 1994, p. 139; Halliday & Matthiessen, 2004, p. 248). Thompson (1996, p. 100) suggests that what distinguishes these material and mental processes is their distinction between "purely mental processes and the outward physical signs of those processes". For example, the process "see" could be analysed as a purely mental perception, whereas "look", "watch" and "stare" express "a conscious physical act involved in perception", and therefore they are categorized into behavioural. The same applies to "hear" (mental) and "listen" (behavioural).

At the boundary between mental and relational are verbal processes which are concerned with saying. Different from 'mental' clauses, the 'verbal' clauses do not require a conscious participant. The main participant is the Sayer, and other participants may include a "Recipient" and "Verbiage" (Halliday & Matthiessen, 2004, p. 255). To some extent, 'verbal' clauses are like 'behavioural' ones, which exhibits certain characteristics of 'material' and/or 'relational' processes. In spite of this, 'verbal' process clauses do display distinctive patterns of their own, the uniqueness of which is shown in their further participant functions to the Sayer: (1) Receiver: "the one to whom the saying is directed", (2) Verbiage: "this function corresponds to what is said, representing it as a class of thing rather than as a report or quote", (3) Target: "this function construes the entity that is targeted by the process of saying" (Halliday & Matthiessen, 2004, pp. 255-256).

At last, existential processes often express the existence of an entity. What happens within existential processes is that "the speaker is renouncing the opportunity to represent the participant (the Existent) as involved" in any event that happens (Thompson, 1996, p. 101). Thompson (1996, p. 101) argues that existential processes enable the speaker to avoid representing the participant as involved in any of the "goings-on". The Existent is the only participant in existential processes, which makes it normally recognizable. Halliday and Matthiessen (2004, p. 257) also comments that "'existential' clauses are not, overall, very

common in discourse … they make an important, specialized contribution to various kinds of text. For example, in narrative, they serve to introduce central participants in the Placement (Setting, Orientation) stage at the beginning of a story (cf. Hasan, 1984: Ch.3)".

Based on what has been described above, transitivity is assumed to reveal a speaker's experience, no matter how they are involved in the experience, or to what extent they associate themselves with different processes. In the exploration of the transitivity system, a complementary perspective is to look into the most common lexical verbs. From this joint perspective, the typical communicative purposes of in particular register can be revealed. What will be referred to in the current research is a very revolutionary and comprehensive study of contemporary English based on a 40 million-word Longman Spoken and Written English corpus (LSWE) including four text types - conversation, fiction, news reportage, and academic prose (Biber et al., 1999). In this study, it has been found that there are important differences in the register distribution of the most common verbs across semantic domains. Among the most frequent 12 lexical verbs, six are activity verbs (get, go, make, come, take, give), five are mental verbs (know, think, see, want, mean), and one is a communication verb (say). The verb 'say' is the most common verb overall in the LSWE Corpus and the only verb to be extremely common in more than one register. Besides, the verb 'get' is the most common lexical verb in any conversational register, being used with a wide range of meanings and grammatical patterns. The verb 'give' is one of the few lexical verbs that is relatively common in all registers; however, this verb is used with activity meanings in most registers, and it typically expresses relational or existence meanings in academic prose.

> "It is noteworthy that two of the most common activity verbs, i.e., get and make, often occur marking causation rather than physical activities. Mental verbs, especially 'know', 'think', 'see', 'want', and 'mean', are particularly common in conversation. These verbs report various states of awareness, certainty, perception, and desire. They typically occur with 'I' or 'you' as subject, and not infrequently occur together in the same utterance. While the mental verb 'see' is also relatively common in academic prose, where it is used to report scientific observations, or to refer readers to related work in the same or in other publication" (Biber et al., 1999, pp. 274-278).

From a theoretical perspective, these resources for the representation of participants can be viewed as constituting a cline along two dimensions: a syntactic-semantic one, congruence, and a pragmatic one, negotiability. Congruence (Halliday, 1994, pp. 341-344) refers to “the typical or unmarked selection of patterns of wordings for the expression of a given meaning”. While negotiability (Martin, 1992) expresses the interpersonal meaning of discourse semantics. Encompassing two dimensions along a cline, different elements in this cline are involved, and there is a complex correlation between them. A general tendency is the less congruent the construction, the more semantic distance between the surface representation and the meaning encoded (Rutherford, 2014). Despite of this tension, it can be confirmed that the interpersonal function is realized through various linguistic elements, and the verb plays an essential role in the manifestation of the interpersonal function. This means the ideational function and the interpersonal function operate a continuous rather than a single dimension. In the semiotic system where the three metafunctions assume the system of grammar, the mechanism for various functions is then able to be integrated in a text (Halliday & Webster, 2006), and vice versa. In a text, those integrated functions can be viewed from different perspectives, and discomposed into different elements.

2.3.3.2 Features related to tenor and the interpersonal function: personal pronouns and modal verbs

As mentioned in Section 2.3.2.2, the tenor of discourse signals the speech roles, i.e., the participants and their relationship, which relates to “who is taking part, to the nature of the participants, their statuses and roles: what kinds of role relationship obtain among the participants, […] and the whole cluster of socially significant relationships in which they are involved” (Halliday & Hasan, 1989, p. 12). Both of tenor and the speech roles are not themselves linguistic features; however, they are among “the contextual category correlating groupings of linguistic features with recurrent situational features” (Gregory & Carroll, 1978, p. 4). In Halliday’s viewpoints, tenor focuses on the selection of interpersonal options, those in the systems of mood, modality, evaluation and modal commitment etc., which are determined by the role relationships in the situational context. In the following ways, interpersonal

relationships influence ‘tenor’: they are determined by the user of the language in a particular situation context, especially by the different formality scale of relationship ranging from intimate, to a familiar or estranged relationship. What have been adopted and analysed in the current research are two language features: personal pronouns and modal verbs, and a more detailed elaboration to them is presented below.

In SFL, the person system is regarded as an important device to realize the interpersonal meaning (e.g., Halliday, 1994; Halliday & Matthiessen, 2004). As claimed in Halliday and Matthiessen (2004, p. 328), “Interpersonal meanings are embodied in the person system, both as pronouns (person as Thing, e.g. she, you) and as possessive determiners (person as Deictic, e.g. her, your)”. From the perspective of their interpersonal role relationship, these pronouns and possessive determiners can be classified into three kinds: the first person pronouns *I*, *we*, *me* and *us* which include the speaker or the writer; the second person pronouns *you* including addressee or the reader; and the third person pronouns *he, she, it, they, him, her* and *them* who are neither the speaker nor the addressee. The corresponding possessives are *my, our, your*, *his*, *her*, *its* and *their*. The first and the second person pronouns are the speech roles in a speech event, while the third person pronouns are the other things except participants in a speech event (Brown & Levinson, 1987). As a composite of the dexis reference system, personal pronouns offer two contexts which provide information. One is linguistic context in a speech; the other is situational context which consists of the environment used directly or the hearer's understandings about any external information, its features and presuppositions, etc. In other words, personal pronouns suggest the relationship between a word and what it refers to in the real world.

Besides, the language feature of modal verbs is also among those concrete ones (e.g., polarity, mood and personal pronouns) that can reveal the important aspects of the interpersonal meaning, such as power or solidarity of the relationship, their social status and roles, as well as their attitudes or opinions. Theoretically, the modal verb (can, may, must etc.) is a modal marker of modality which is also concerned with mood (subjunctive etc.), can be treated as a single grammatical category, as well as a semantic category which operates at the sentence

level but has effects on the whole text (e.g., Halliday, 1994).

Modality refers to “the area of meaning that lies between yes and no — the intermediate ground between positive and negative polarity” (Halliday & Matthiessen, 2004, p. 618). More specifically, modality can be divided into modalization and modulation according to the distinction between proposition and proposal. When it comes to modalization, it refers to speaker’s assessments of probability (possible, probable, certain) and usuality (sometimes, usually, always); as for modulation, it means the speaker's assessment determined) of obligation (allowed, supposed, required) and inclination (willing, keen, determined), as shown in Figure 2.7:

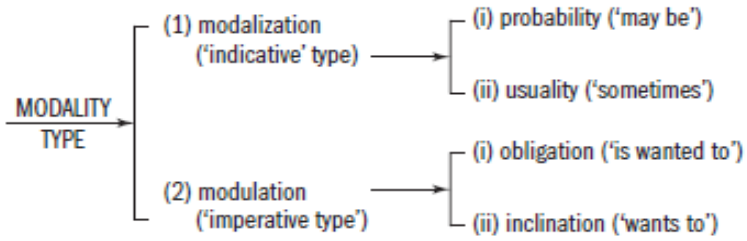


Figure 2.7: System of types of modality (Halliday & Matthiessen, 2004, p. 618)

This system shows that, in the process of communication, a speaker may express a higher or lower degree of certainty on what has been proposed or put a higher or lower pressure/command on the hearer. In order to analyse the degree and scales of the modality, Halliday and Matthiessen (2004, p. 620) have set up three basic values that are attached to the modal judgment: “high” (must, ought to, need, has to, is to), “median” (will, would, shall, should) and “low” (may, might, can, could). In the current research, however, the analysis is based on three categories of modal verbs from Biber (1988, p. 241): Possibility modals (can, may, might, could), Necessity modals (ought, should, must), Predictive modals (will, would, shall). This categorization is based on their functions in language: Possibility modals are for linguistically realizing different perspectives on a topic, necessity modals are for expressing persuasiveness, and predicative modals are for referring to the future.

2.3.3.3 Features related to mode and the textual metafunction: lexical density and grammatical complexity

As mentioned in Section 2.3.2.3, the mode of discourse mode focuses on "what part the language is playing, what it is that the participants are expecting the language to do from them in that situation: […] what is being achieved by the text in terms of such categories as persuasive, expository, didactic, and the like" (Halliday & Hasan, 1989, p. 12). Accordingly, determined by the symbolic organization taken by the interaction, its function in the context, etc., mode of discourse functions as a resource for building interaction in communication and mediates the role that plays in language with two dimensions, namely, ideational and interpersonal. As stated in Halliday (1978, p. 130), "language can effectively express ideational and interpersonal meanings only because it can create text", thus we can call textual function the "enabling" function in comparison with the other two. The mode of discourse is highly sensitive to medium (e.g., Neumann, 2010), as well as the positioning between the extremes of spoken and written languages (e.g., Eggins, 2004), and there are many language features characterizing these. What have been selected for the current research are the features of lexical density and grammatical complexity, the detailed elaboration of which is presented below.

Both lexical density (LD) and grammatical complexity (GC) are concerned with the intricacy of the text. As stated in Halliday and Matthiessen (2004, p. 654): "typically, written language becomes complex by being lexically dense: it packs a large number of lexical items into each clause; whereas spoken language becomes complex by being grammatically intricate: it builds up elaborate clause complexes out of parataxis and hypotaxis". This claim has an implication: whether or not the text tends to be spoken or written can be identified through the exploration of LD and GC. Besides, LD and GC also found to be related to certain registers. In some empirical studies, such as Schleppegrell's (2002) work, it has been found that academic texts are complex not only lexically but also grammatically. Based on a quantification of the amount of cognitive processing to decode school based texts, the complexity of academic texts is characterized by the use of longer words, varied vocabulary and lexical density, and more

complex grammar with the use of more subordinate clauses and passive forms, etc.

In various studies which focus on the measurement of LD or GC, different approaches have been proposed and employed, accompanied by the emergence of different definitions. Both of Halliday's (1996) and Ure's (1971) works on LD have been influential, which have focused on the balance between lexical 'content words' and grammatical 'function words'. Based on the distinctions of words with and without lexical properties (Ure, 1971), Ure and Ellis (1977, p. 207) defines lexical density as "the proportion of words carrying lexical values (members of open-ended sets) to the words with grammatical values (items representing terms in closed sets). LD percentage, according to Ure (1971), generally tends to be over 40% in written texts and less than 40% in the spoken ones. By contrasting written and spoken versions of one and the same text, Eggins (1994, p. 61) also supported Ure's argument: according to the findings of her study, LD was 9% higher in favour of the written text (33% vs. 42%). Later on, Halliday (1985) provided a further development to the concept of LD, as well as pointing out the importance of discriminating between lexical items and grammatical items. According to Halliday (1987, p. 60),

> "Lexical density is the proportion of lexical items (content words) to the total discourse. It can be measured in various ways: the ratio of the lexical items either to the total running words or to some higher grammatical unit, most obviously the clause".

Based on this definition, two formulas can be made, as presented below:

1) LD = Number of lexical words/Total running words (Tokens)×100 %
2) LD = Total number of lexical words/Total Number of ranking clauses×100 %

The definition by Eggins (2004) is very close to the first formula presented above: LD expresses "the number of content carrying words in a text/ sentence [nouns, main part of verb, adverbs, adjectives] as a proportion of all the words in the text/ sentence" (Eggins, 2004, p. 97). As for the lexical and grammatical items, lexical items are often called content words or open class words due to the possibility to easily include new members of the class; however, grammatical items are called closed classes, since new prepositions or pronouns seldom enter the language (Johansson, 2008).

In the current research, I will adopt Halliday's (1996, p. 347) measurement, which calculates "the proportion of content (lexical) words - basically nouns, full verbs, adjectives, and adverbs derived from adjectives - over the total number of words in a text". This approach for the finding of lexical item is more practical. By using this approach, a continuum between the lexis (content words) and the grammar (non-content words) of English can be established.

Apart from LD, GC is another language feature to be explored. In GC, what 'complexity' really means and how it can be measured remains unresolved in linguistics. Up to now, there are few studies which provide any widely accepted formulation based on the notion of complexity. In the current research, fourteen syntactic/grammatical complexity measures, which were originally selected for an automatic analysis of syntactic complexity in second language writing, will be taken into account, and analysed for an understanding of the difference between spoken and written language (Lu, 2010). The fourteen grammatical complexity measures that are adopted in the current research were selected from Wolfe-Quintero, Inagaki, and Kim (1998) and Ortega (2003). The detailed measures are presented in Table 2.2. These GC features/measures are both related to T-unit-based measures and the clausal subordination, which have been demonstrated to be the best complexity measures so far by Wolfe-Quintero et al. (1998). Experimental results show that this system using fourteen measures achieves very high reliability on the data from the corpus (Lu, 2010).

Table 2.2: The fourteen syntactic/grammatical complexity measures automated (Lu, 2010, p. 479)

Grammatical complexity features /measures	Code	Definition
Type 1: Length of production unit		
Mean length of clause	MLC	# of words / # of clauses
Mean length of sentence	MLS	# of words / # of sentences
Mean length of T-unit	MLT	# of words / # of T-units
Type 2: *Sentence complexity*		
Sentence complexity ratio	C/S	# of clauses / # of sentences
Type 3: *Subordination*		
T-unit complexity ratio	C/T	# of clauses / # of T-units
Complex T-unit ratio	CT/T	# of complex T-units / # of T-units
Dependent clause ratio	DC/C	# of dependent clauses / # of clauses
Dependent clauses per T-unit	DC/T	# of dependent clauses / # of T-units
Type 4: *Coordination*		

Coordinate phrases per clause	CP/C	# of coordinate phrases / # of clauses
Coordinate phrases per T-unit	CP/T	# of coordinate phrases / # of T-units
Sentence coordination ratio	T/S	# of T-units / # of sentences
Type 5: Particular structures		
Complex nominals per clause	CN/C	# of complex nominals / # of clauses
Complex nominals per T-unit	CN/T	# of complex nominals / # of T-units
Verb phrases per T-unit	VP/T	# of verb phrases / # of T-units

As can be seen in this table, all the GC measures can be categorized into the following five types: The first type consists of three measures that concern the length of production at the T-unit level, namely, mean length of clause (MLC), mean length of sentence (MLS), and mean length of T-unit (MLT). The second type consists of a sentence complexity ratio (clauses per sentence, or C/S). The third type comprises four ratios that reflect the amount of subordination, including a T-unit complexity ratio (clauses per T-unit, or C/T), a complex T-unit ratio (complex T-units per T-unit, or CT/T), a dependent clause ratio (dependent clauses per clause, or DC/C), and dependent clauses per T-unit (DC/T). The fourth type is made up of three ratios that measure the amount of coordination, namely, coordinate phrases per clause (CP/C), coordinate phrases per T-unit (CP/T), and a sentence coordination ratio (T-units per sentence, or T/S). The fifth and final type consists of three ratios that consider the relationship between particular syntactic structures and larger production units, i.e. complex nominals per clause (CN/C), complex nominals per T-unit (CN/T), and verb phrases per T-unit (VP/T).

All these measures are T-unit based. T-unit analysis has been used widely for the measurement of the overall syntactic/grammatical complexity of speech and writing samples, both for first language research and, more recently, for research in second language learning/teaching. T- unit was coined by Hunt in 1965, and defined as “a main clause plus all subordinate clauses and noncausal structures” (Hunt, 1965, p. 360) or “the minimal syntactic unit containing one independent clause and the dependent clauses (if any) syntactically related to it” (O'Donnell, 1974, p. 103). In different studies, the use of this unit has been varied as for the consideration of embedded clauses as dependent clauses and coordinated clauses as one or two different units (Foster, Tonkyn, & Wigglesworth, 2000, p. 360). A compatible approach to T-unit is the clause complex proposed by Halliday (Thompson, 1996). The

similarities between them are both of them include a main clause and all the clauses that relate to it. And both of them do not address sentence fragments. However, they also differ from each other in that Halliday (1985) considers coordinated clauses within the unit, which reflects the process of speaking as a reflection of how individuals present and relate events in the world.

Although there is not a commonly accepted approach to explore GC, it is a question worthy to be answered and investigated. From both empirical and theoretical perspectives, it has been well established that speech and writing are grammatically complex in dramatically different ways. More importantly, it is well established that the grammatical features that are more closely associated with clausal subordination features are actually much more common in conversation than in academic writing. Thus, if we focused on only clausal subordination features, we would inevitably narrow our vision and conclude that conversational discourse is more complex than academic writing. In contrast, the complexities of academic writing are phrasal rather than clausal (Biber, Gray, & Poonpon, 2011). To solve this problem, the fourteen syntactic/grammatical complexity measures are adopted and applied for the current research which provides a more specific grammatical characteristic than the traditional T-Unit-Based Measures of Grammatical Complexity proposed by Hunt (1965).

In all, the relationship of three variables of register theory, field, tenor, and mode of discourse are not separated from each other, but closely connected and make the effect as a whole. The three variables are correspondingly in relation to ideational meaning, interpersonal meaning and textual meaning, i.e., meanings in tenor are realized through interpersonal metafunctional choices, field meanings are realized through ideational metafunctional choices, and textual meanings are realized through mode metafunctional choices.

2.3.4 Register variation and Biber's multi-dimensional analysis

2.3.4.1 Overview of the multi-dimensional approach to register variation

As indicated in the previous section, register is a cover term for any language variety defined by its situational characteristics, including the speaker's purpose, the relationship between

speaker and hearer, and the circumstances. Some researchers, such as Halliday (1988) and Brown and Fraser (1979), emphasize the importance of linguistic co-occurrence theoretically and correlate the notion of register with the issue of 'probability'. On these bases, the emergence and the application of Biber's MD approach makes it possible to move from a theoretical to a practical deliberation. Presented below is a conceptual and methodological overview of this approach, as well as a review of some previous MD studies.

Multi-dimensional analysis is a "methodological approach that applies multivariate statistical techniques (including factor analysis and cluster analysis) to the investigation of register variation in a language" (e.g., Biber, 2004a, p. 15). This approach was originally developed to analyse the full range of spoken and written registers in a language. And later on, an increasing number of studies have been conducted by using this approach, which focus on different domains and other languages. Therefore, it can be said that this approach has been well received in register studies in literature as it establishes a link between form and function. Among these studies, most of them are conducted by Biber (e.g., 1988, 1995, 2004a), and partly by some other researchers, such as Westin and Geisler (2002)，Van Rooy and Terblanche (2009), Xiao (2009), Asencion-Delaney and Collentine (2011), etc. From a methodological perspective, these previous studies use slightly different MD approaches: some limit themselves to factor analysis (e.g., Biber, 1995; Biber et al., 2002), and some carry out a cluster analysis as a further step to the previous factor analysis (e.g., Biber & Kurjian, 2006; Grieve et al., 2011). Most of these studies associate those previously identified dimensions to the ones identified in their own analysis. Methodologically, a full MD study (e.g., Biber, 1995) includes the following steps:

1) An appropriate corpus is designed based on previous research and analysis. Texts are collected and input into the computer.
2) Computer programs are used to identify or 'tag' all relevant linguistic features in the texts, and all texts are edited to ensure that the linguistic features can be accurately identified.
3) Additional computer programs are run to compute normed frequency counts of each linguistic feature in each text of the corpus.
4) The co-occurrence patterns among linguistic features are identified through a factor

analysis of the frequency counts. The ‘factors’ from the factor analysis are interpreted functionally as underlying dimensions of variation.

5) Dimension (the factor loadings) scores for each text with respect to each dimension are computed; the mean dimension scores for each register are then compared to analyse the salient linguistic similarities and differences among the registers.
6) The factors identified can be further used as predictors in a cluster analysis, a multivariate statistical technique to group sub-corpora into clusters based on shared linguistic characteristics. The texts grouped together in a cluster are maximally similar linguistically, while the different clusters are maximally different linguistically.

Following these steps, it is possible to conduct a full multi-dimensional analysis or to apply the dimensions originally identified by Biber to the current research. Among these procedures, cluster analysis is an additional element of the MD approach. In many previous MD studies, some limit their studies to factor analysis (e.g., Biber, 1995; Biber et al., 2002), and some use cluster analysis as well (e.g., Biber & Kurjian, 2006; Grieve et al., 2011). In the current research, considering that the main objective is to identify the dimensions, as well as the CB text types, a complete procedure including both factor analysis and cluster analysis will be followed. What follows is a more detailed introduction to factor analysis and cluster analysis (Biber, 1995).

In a factor analysis, a large number of original variables (in this case the linguistic features) are reduced to a small set of derived, underlying variables quantitatively - the factors (in this case the dimensions). And then, a qualitative analysis is required for the interpretation of the functions associated with each set of co-occurring linguistic features. This is based on the characteristics of the dimensions of variation that have both linguistic and functional content, and accordingly an assumption that linguistic co-occurrence patterns (central to MD analyses in that each dimension represents a different set of co-occurring linguistic features) reflect underlying communicative functions. And then, the factors identified in the factor analysis can be further used as predictors in a cluster analysis, a multivariate statistical technique to group texts into clusters based on shared linguistic characteristics.

2.3.4.2 Summary of the 1988 MD analysis of English registers

The multi-dimensional approach to register variation was originally developed for comparative analyses of spoken and written registers in English (e.g., Biber, 1985). The 1988 multi-dimensional analysis of register variation in English is usually considered as the first major MD study in literature, since it provides the fullest account of multidimensional methodology and a synchronic analysis of the relations among spoken and written registers. This MD analysis exceeds many early spoken/written language studies, in that it views spoken and written language more systematically, involving more complex and continuous parameters rather than merely a simple dichotomy. In this MD approach, it is assumed that all registers have distinctive linguistic characteristics associated with their defining situational characteristics. Thus, it can be said that the Biber's 1988 MD study describes the relations among the full range of spoken registers and the full range of written registers, as well as comparing speech and writing within a complex situational context. His book *Dimensions of register variation: A Cross-linguistic comparison* published in 1995 is a refined work to the study in 1988. Therefore, the review presented below includes some details from Biber's work in 1988 and 1995 as well. As can be seen in Table 2.3, a more detailed list of functions, linguistic features, and characteristic registers for the major English dimensions identified are included.

Table 2.3: Summary of functions, (partial) linguistic features, and characteristic registers for the five major English dimensions identified in Biber (1988)

Dimensions	Functions	Linguistic features	Characteristic registers
Dimension 1 'Involved vs. informational production'	involved/ interpersonal focus; interactive; personal stance	present tense verbs: 0.86	conversations, personal letters, public conversations
		second-person pronouns: 0.86	
		demonstrative pronouns: 0.76	
		first-person pronouns: 0.74	
		pronoun *it*: 0.71	
		wh-questions: 0.52	
		possibility modals: 0.50	

		wh-clauses: 0.47	
	informational production	nouns: - 0.80	informational exposition, e.g., official documents, academic prose
		word length: - 0.58	
		type/token ratio: - 0.54	
Dimension 2 'Narrative vs. non-narrative concerns'	narrative	past tense verbs: 0.90	fiction
		third-person pronouns:0.73	
		perfect aspect verbs: 0.48	
		public verbs: 0.43	
		present participial clauses: 0.39	
	non-narrative	(present tense verbs: - 0.47) (attributive adjectives: - 0.41)	exposition, broadcasts, professional letters, telephone conversations
Dimension 3 Situation-dependent Reference vs. elaborated reference	situation-independent reference	time adverbials: 0.60	official documents, professional letters, written exposition
		place adverbials: 0.49	
		adverbs: 0.46	
	elaborated reference	*wh* relative clauses on object positions: - 0.63	broadcasts, conversations, fiction, personal letters
		wh relative clauses on subject positions: - 0.45	
		nominalizations: - 0.36	
Dimension 4 Overt expression of persuasion	overt expression of persuasion or argumentation	Infinitives: 0.76	professional letters, editorials
		prediction modals: 0.54 conditional subordination: 0.47	
		necessity modals: 0.46	
		possibility modals: 0.37	

	not overtly persuasion	-	broadcasts, press reviews
Dimension 5 Abstract vs. Non-abstract Style	abstract style	-	technical prose, other academic prose, official documents
	non-abstract style	conjuncts: - 0.48	conversations, fiction, personal letters, public speeches, public conversations, broadcasts
		agentless passives: - 0.43	
		past participial adverbial clauses: - 0.42	
		by-passives: - 0.41	
		past participial postnominal clauses: - 0.40	

As shown in this table, five major dimensions of variation in a wide range of spoken and written registers are identified. Each dimension comprises a set of co-occurring linguistic features; each includes the configuration of different spoken and written registers, and their distinct functional labels. On a dimension/factor, the loadings of the co-occurring linguistic features are greater than +/-0.35, which were considered important enough to be used in the computation of dimension scores (cf. Chapter 4). In Biber (1988), 67 linguistic features across 481 spoken and written texts of contemporary British English were selected, and the dimensions identified represent their co-occurrence distributions. The texts, taken from the Lancaster-Oslo-Bergen Corpus and the London-Lund Corpus, represent 23 major register categories (e.g., academic prose, press reportage, fiction, letters, conversations, interviews, radio broadcasts, public speeches). Most of the dimensions identified consist of two groupings of features, i.e., two sets of features with positive and negative loadings. “The positive and negative sets represent features that occur in a complementary pattern. That is, when the features in one group occur together frequently in a text, the features in the other group are markedly less frequent in that text, and vice versa” (Biber et al., 2002, p. 14).

Among the dimensions identified, there are some relationships between them. Overall, three dimensions identified are assumed to have a systematic relationship with spoken and written registers (Dimension 1: Involved versus Informational Production; Dimension 3: Situation-dependent vs. Elaborated Reference; Dimension 5: Abstract versus Non-abstract Style), because many co-occurring language features from these dimensions are typical features characterizing spoken and written registers. In contrast, the other two dimensions cannot be sharply distinguished along spoken and written languages (Dimension 2: Narrative Discourse; and Dimension 4: Persuasion).

Dimension 1, labelled as 'Involved versus informational production', marks highly involved, interactional text production versus high density and highly informational text production. Among the linguistic features that are representative of involved production are first and second person pronouns, demonstrative pronouns, possibility modals, etc. The linguistic features that are representative of informational production are nouns, word length, type-token ratio, etc. (Biber, 1988). Both word length and type-token ratio mark a high density of information, with a higher lexical diversity and the use of longer words which tend to be rarer and more specific in meaning than shorter words (Zipf, 1949). For many researchers (e.g., Ravid & Berman, 2006), the information density in spoken versus written registers has always been an interesting issue to be explored.

Dimension 3, labelled as 'Elaborated versus situation-dependent reference', characterizes highly context-independent elaborated reference versus situation-dependent reference. Among the linguistic features that are representative of situation-dependent reference are time and place adverbials, which are usually used for locative and temporal reference (e.g., above, behind; earlier, soon). The linguistic features representative of elaborated reference are WH-relative clauses and nominalizations, which specify the identity of referents in an elaborated manner, and therefore leave no doubt about the intended referent. As for the relationship of this dimension to spoken/written language, it has been found that written language is typically less dependent on the spatial and temporal situation in which it is produced than spoken language (e.g., Cook-Gumperz, 2006).

Finally, Dimension 5, labelled ‘Abstract versus non-abstract style’, characterizes discourse which is typically abstract, technical and formal in content. The co-occurrence of agentless passives, by-passives, and conjuncts, etc. in this dimension indicates that the texts with high factor 5 scores are passive forms, presenting information with reduced emphasis on the agent, and giving prominence to the ‘patient’, the entity acted upon. In Biber (1988), academic prose and official documents are the two registers that are most marked by the use of Dimension 5 features. Written language is often more ambiguous or obscure, characterized by the use of passive constructions (Halliday, 1987). In some particular contexts, such as the school context, written language used plays a very important role in making meaning, so their uses have to be expanded beyond those resources needed for everyday interaction, by means of taking on the more technical and abstract meanings necessary for construing academic knowledge (Schleppegrell, 2004).

The results of Biber’s 1988 study also show that the various register varieties included in the corpus have different scores on the dimensions identified. Official documents, for instance, has one of the lowest scores in Dimension 1, showing its typical informational characteristics. However, face-to-face conversations have one of the highest scores and are thus highly involved (Figure 2.8).

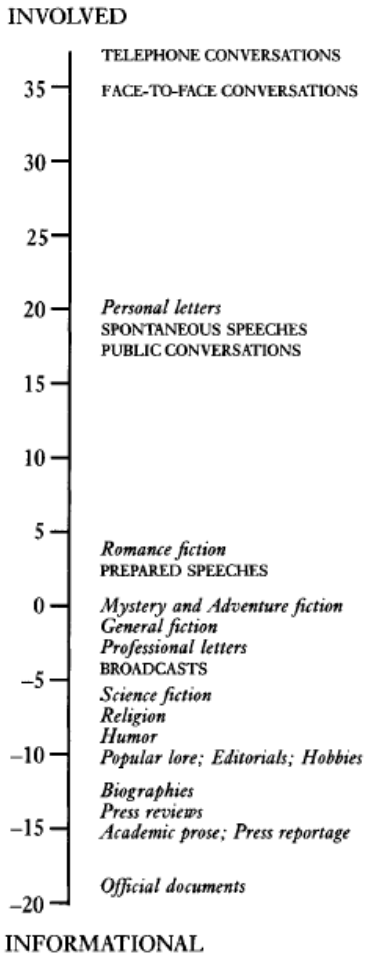


Figure 2.8: Mean scores of English Dimension 1 for twenty-three registers: 'Involved versus Informational Production' (Biber, 1995, p. 146)

In Dimensions 3 and 5 (cf. Figure 2.9 & Figure 2.10), some other distinctive registers are able to be found as well: news broadcasts and conversations, which contain different pervasive linguistic features. Frequently co-occurring language features in conversations include first and second person pronouns, demonstrative pronouns, adverbs, etc., while frequent features in news are nouns, prepositions, adjectives and high type-token ratio. Besides, it was also found that the frequent features of conversations are relatively infrequent in news, and the frequent features of news are relatively infrequent in conversations.

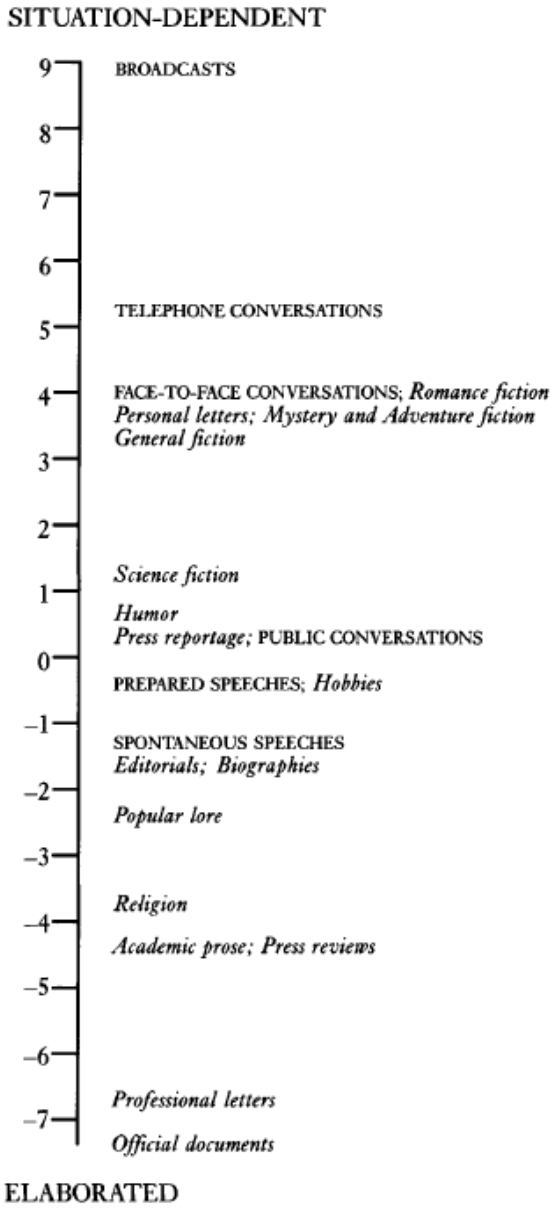


Figure 2.9: Mean scores of English Dimension 3 for twenty-three registers: 'Situation-dependent versus Elaborated Reference' (Biber, 1995, p. 157)

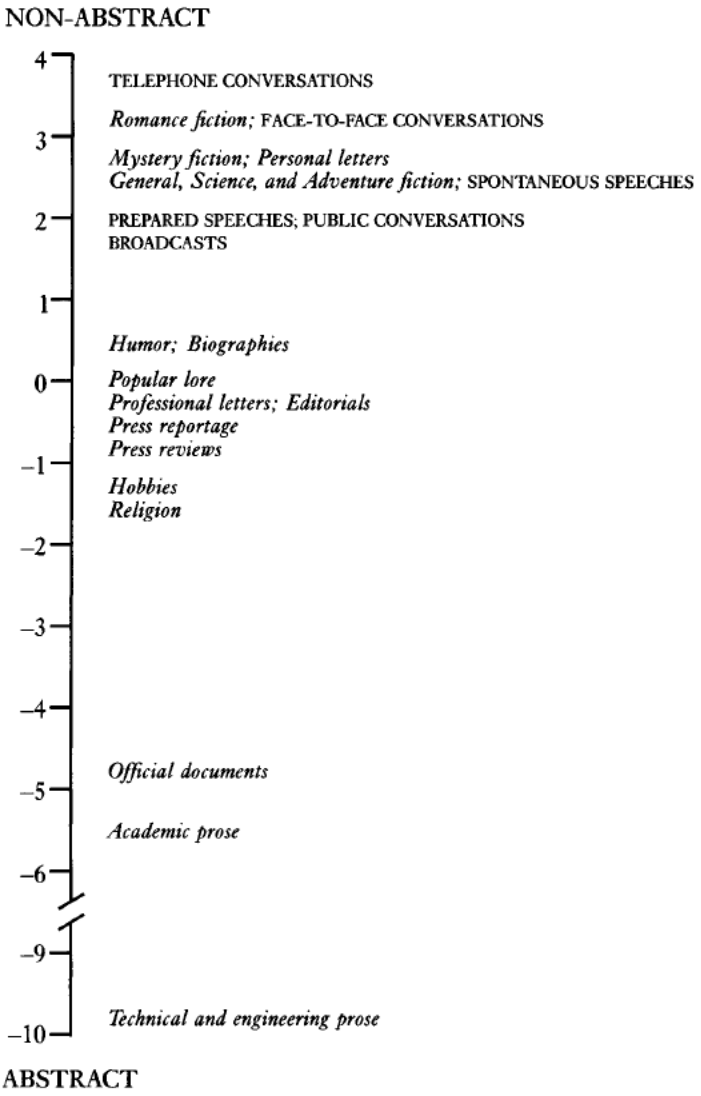


Figure 2.10: Mean scores of English Dimension 5 for twenty-four registers: 'Non-abstract versus Abstract style' (Biber, 1995, p. 165)

Among the five dimensions identified, the other two dimensions, i.e., Dimension 2: Narrative Discourse and Dimension 4: Persuasion, have no systematic relationship to spoken and written languages. Dimension 2, which is labelled as Narrative vs. Non-narrative discourse, includes some language features associated with past time narration on its positive loading: past tense verbs, third-person pronouns, perfect aspect verbs, public verbs, etc. Among these language features, past tense verbs and third-person pronouns are markers of narration. They describe the major past events that constitute the backbone of the narrative, which are often used to refer to the animate participants in the narrative. Public verbs are used to report the direct and indirect speech acts of these participants, including verbs such as *assert*, *declare*, *report, say,* and *tell*. Thompson (1983) claims that these forms are able to create vivid imagery

in interactive discourse, and their co-occurrence with narrative features apparently reflects their use for describing something in narratives.

Dimension 4 which is labelled as 'Overt Expression of Persuasion' in Biber (1988) is later renamed as 'Overt Expression of Argumentation' in Biber (1995). This does not mean Dimension 4 in Biber (1988) is fundamentally distinct from that in Biber (1995). The rename of this dimension is possibly owing to a consideration that "there is no general characterization as persuasive or not; rather certain texts within these genres are persuasive, while others are not" (Biber, 1988, p. 151). In this dimension, most registers are unmarked, apart from professional letters and editorials, which means these two registers have some notable and co-occurring language features: three modal classes, together with conditional subordination distinguish persuasion or argumentation. With a dominant use of these language features, different positions by authors/ speakers and the circumstances under which other actions or events may occur are exhibited clearly.

2.3.4.3 Application of MD approach in other studies

Since the groundbreaking work by Biber (1988), the MD approach has been widely used for the investigation of register variation. From different perspectives, the focus has been put on traditional media types (Biber, 2006; Biber et al., 2002), or some recently emerging web-based social media types (Biber & Kurjian, 2006; Grieve et al., 2011); the focus has also been put on the study of synchronic linguistic variation (e.g., Biber, 1994, 1995; Finegan & Biber, 1994) as well as of diachronic variation (Atkinson, 1992; Biber & Finegan, 1989a, 1992, 2001; Geisler, 2002, 2003; Westin & Geisler, 2002), or the register variation on monolingual language (Asencion-Delaney & Collentine, 2011; Biber, 1988; Biber & Hared, 1992), and multiple languages (e.g., Biber, 1995; Xiao, 2009).

Among various research studies using MD approach, Biber and his colleagues have always been the pioneers, doing extensive as well as intensive research works. Based on the groundbreaking work in 1988, Biber's study in 1995 extended and refined the earlier 1988 MD study, exploring register variation across English language as we'll as languages other than

English, i.e., Somali, Korean and Nukulaelae Tuvaluan. In early times, traditional media types in English language or across languages are more possible to be chosen as the target language analysed. For example, Biber (2001) identified six dimensions of variation for 18th century written and speech-based registers; and then, Biber et al. (2002) identified the dimensions operating among the range of spoken and written registers at U.S. universities by using the TOEFL 2000 Spoken and Written Academic Language Corpus. More recently, Biber, Davies, Jones, and Tracy-Ventura (2006) provided the first multi-dimensional analysis of native speaker Spanish on the basis of a 20-million-word Spanish corpus with written and oral data representing 19 registers. The common ground shared by these studies is that a full range of spoken and written registers are included in the data, allowing for the linking between existing registers and salient language features in a dimension. In comparison with this, many MD studies conducted by other researchers have focused on specialized registers. However, these specialized registers are not restricted to a specific domain; rather, they are from a wide range of discourse domains in both spoken and written languages in English, or in other languages. For example, Grabe (1987) has investigated the dimensions of variation for written expository registers. Connor-Linton (1989) identified dimensions of variation in the conversational styles of Soviet and American participants in 'Spacebridge' interactions on TV talk shows. Reppen (2001) compared the MD characteristics of elementary school spoken and written registers. Conrad (2001) compared research articles, textbooks, and student writing in two academic disciplines: biology and history. Asencion-Delaney and Collentine (2011) conducted an MD analysis of a corpus of L2 Spanish writing generated by university-level learners. And SARDINHA, KAUFFMAN, and ACUNZO presented an MD analysis of different spoken registers (ranging from face-to-face conversation to radio, TV shows, and telephone conversation, among others) in Brazilian Portuguese.

More recently, with the emergence of internet-based social media types, i.e., different CMC forms, such as emails, blogs, chatrooms, etc., Biber and his colleagues have also shifted their focus from traditional registers/media types to these new language phenomena. Two MD studies have been conducted on CMC: MD analysis of web registers by Biber and Kurjian (2006), and on blogs by Grieve et al. (2011). Given the focus on corporate blogs in the current

research which is especially relevant to the two MD studies, they will be reviewed in detail in the following part.

Methodologically, both studies have employed a full MD analysis including both factor analysis and cluster analysis. In Biber and Kurjian (2006), the web documents that represent two Google categories, i.e., 'Home' and 'Science', were investigated. In this study, a 3.7 million-word corpus is constructed by a stratified sample of web pages, representing the major sub-categories within each of these two top-level categories (e.g. 'apartment living', 'consumer information', and 'cooking' within the 'Home' category). In this study, it was decided that four factors can be selected as optimal. Their interpretation is presented below: Dimension 1 can be labelled as 'Personal, involved narration', since this dimension combines some of the major linguistic features and functions associated with the first two dimensions in the Biber (1988) study. Some linguistic features in this dimension reflect personal involvement and the expression of personal stance (e.g. 1st person pronouns, certainty/mental verb + *that*-clause, certainty adverbials, likelihood/mental verb + *that*-clause), and these features co-occur with some features used for narration, such as past tense verbs, 3rd person pronouns, perfect aspect verbs, and communication verb + *that* clause. These co-occurring functions (personal involvement/interaction and narrative discourse) in Biber and Kurjian (2006) are actually the two distinct dimensions in the 1988 MD study of spoken and written registers, which include many interactive conversations (face-to-face and telephone), as well as some excerpts from novels and fictional short stories, relying heavily on stereotypical narrative features. Besides, Dimension 2 can be labelled as *Persuasive/argumentative discourse*. This dimension includes many of the defining features (conditional adverbial clauses, possibility modals, necessity modals, prediction modals, etc.) grouped on Dimension 4 in the 1988 MD analysis. In terms of Dimension 3, it can be labelled as 'Addressee-focused discourse', characterized by a combination of features (2nd person pronouns, progressive aspect verbs, desire verb + *to*-clause, activity verbs, etc.) functioning as advice or service giving in web documents. Finally, Dimension 4 can be interpreted as 'Abstract/technical discourse', in that this dimension consists entirely of features (nominalizations, abstract nouns, long words, attributive adjectives, etc.) for presenting some complex and technical information. This set of features is very similar

to the negative co-occurring features on Dimension 1 in the 1988 MD analysis. Besides, this MD analysis also identified and interpreted some web 'text types'. Totally, there are 8 clusters identified; among these clusters, clusters 1-4 have relatively intermediate dimension scores, while clusters 5-8 have notably large positive or negative dimension scores, as well as marked characterizations. Thus, clusters 5-8 can be assigned functional labels, reflecting their interpretations as web 'text types': Cluster 5: 'Informational Discourse'; Cluster 6: 'Persuasive Advice'; Cluster 7: 'Technical Discourse'; Cluster 8: 'Personal Narrative'.

Another MD study that is closely related to the current research explores register variation in blogs (Grieve et al., 2011). The corpus compiled for this study contains approximately 2 million words and represents the modern (2003 - 2005) American blog variety of the English language. At first, in the factor analysis, four principal dimensions of linguistic variation were identified: the informational vs. personal focus dimension, the addressee focus dimension, the thematic variation dimension, and the narrative style dimension. The identification of these dimensions depends on the function of the linguistic features loading on each factor/dimension, as well as the samples prototypical of a particular factor/dimension. Particularly, Dimension 1 can be labelled as 'informational vs. personal focus', since most of the positive features on Dimension 1, e.g., prepositions, attributive adjectives, nominalizations, passives, and various post-nominal modifying clauses, are associated with nouns and noun modification. Thus the functional significance of this grouping of features in actual texts tends to be an information-focused style. In contrast, many language features with high negative loadings are verbal, and involved, e.g. first person pronouns, time adverbials, place adverbials, progressive verbs, to clauses with desire/intent/decision verbs, quantity nouns, activity verbs, while some other language features are associated with a narrative style, e.g., past tense, activity verbs, progressive aspect, and one feature is associated with personal plans or desires, i.e., *to* clauses with desire/intent/decision verbs. Considering that texts with high negative scores on Factor 1 have one major topic, i.e., their author, the texts with high loadings seem to both recount and comment on the lives of their authors.

Dimension 2 can be labelled as the 'Addressee Focus Dimension', since this factor seems to

reflect functional variation in blogs associated with the degree of interaction and the focus on the addressee: blogs. In particular, many positive features (e.g. present tense, second person pronouns, demonstrative pronouns, indefinite pronouns, WH questions, possibility modals, predictive modals, conditional subordination, necessity modals, mental verbs) on Dimension 2 are associated with interactive discourse and addressee focus. For example, the high frequency of second person pronoun *you* is most obviously associated with an interactive style; the high frequency of other pronouns indicates an interactive style because the frequent use of pronouns implies that the reader is in the same basic frame of reference as the author. Similarly, WH-questions are directly related to mutual communication and interaction, and other features, such as modals, mental verbs, and conditional subordination, are used in blogs for giving advice or instruction to the reader. This identification also relies on some sample texts prototypical of a particular factor: blogs scoring highly positively on Dimension 2 directly refer to and interact with their readership, whereas blogs scoring highly negatively on Dimension 2 essentially ignore their audience, choosing instead to simply express information.

Dimension 3 is named as 'Thematic Variation', because many positive loading features (e.g. demonstrative pronouns, pronoun *it*, predicative adjectives, clausal coordination, conjuncts, predicative adjectives, factive adverbs, likelihood adverbs) appear to be associated with a spoken and conversational style, and many blogs that have large positive scores on Factor 3 shift topic rapidly, as is common in spoken discourse. Finally, Dimension 4 is labelled as 'Narrative Style', since most of the features (e.g. past tense, third person pronouns, *that* deletion, factive verbs with *that* clauses, certain forms of adverbial subordination, *to* clauses with speech acts verbs, and *to* clauses with modality/cause/effort verbs) that load positively on Dimension 4 are associated with a narrative style. Among these language features, past tense verbs and third person pronouns are clearly the markers of narration. On the basis of the four dimensions identified, two major blog text types were identified by the cluster analysis: personal blogs and thematic blogs. It is commonly shared by the two text types that both of these blog types are characterized by a highly personal and conversational style, which appears to be the standard blog voice. They also distinguish with each other in terms of the content of the blogs: one focuses on their authors' lives, and the other focuses on impersonal

or informational topics. Besides, a third blog type can be identified: a bog type which is basically informational, like the thematic blog, but is written in a distinctly non-blog-like style, which is similar to standard informational writing. Considering that this blog type appears to be quite rare, it is assumed that there are only two basic types in blogs: personal blogs and thematic blogs (Grieve et al., 2011).

These previous MD studies, particularly the two relevant studies (Biber & Kurjian, 2006; Grieve et al., 2011) and the large-scale groundbreaking work by Biber (1988, 1995), illuminate the current MD study on corporate blogs in different ways. Overall, it seeks to reveal what dimensions and text types might be included in corporate blogs, and how corporate blogs distinguish web registers, and general blogs in terms of the diverse text types and dimensions identified. At the same time, it also aims to understand how the corporate blog is positioned between spoken and written languages.

In the case of diverging functional dimensions identified in these previous MD studies, several issues concerning the similarities and differences between these dimensions may arise: Firstly, a dimension identified in one register may not be considered register-specific. Actually, in almost all these previous studies, the first dimension is associated with an informational focus versus a personal focus. It is perhaps not surprising that Dimension 1 in the original 1988 MD analysis was strongly associated with an informational versus (inter) personal focus, given that the corpus in that study ranged from spoken conversational texts to written expository texts. However, it was unexpected that a similar spoken/written dimension realized by essentially the same set of co-occurring linguistic features is also fundamentally important in online discourse and some highly restricted discourse domains, like blogs. Secondly, there is possibly an overlap of communicative functions in some dimensions identified in different studies, which are usually based on different language data, representing different discourse domains. Therefore, in any of these dimensions, some overlapping communicative functions and some additional special characteristics are included. For example, the second dimensions found in most MD analyses (e.g., Biber, 2001, 2004a; Biber, 2006) corresponds to narrative discourse, reflected by the co-occurrence of features like past tense, 3rd person pronouns, perfect aspect,

and communication verbs. Apart from this, some other sets of language features are contained in this dimension as well. In a more restricted register domain, such as the job interview corpus (White, 1994), the narrative dimension reflected a fundamental opposition between personal/specific past events and experiences (past tense verbs co-occurring with 1st person singular pronouns) versus general practice and expectations (present tense verbs co-occurring with 1st person plural pronouns). As claimed by Biber (1995, p. 135), "it is not the case that MD analyses assume that there is a single communicative function associated with each linguistic feature. Rather, in many cases, features can have somewhat differing functions in different kinds of text; further, features can have functions defined at different levels of generality". Thirdly, some dimensions that are unique to the particular discourse domain have also been identified. For example, the factor analysis in Reppen (2001) identified a dimension of 'Other-directed idea justification' in elementary student registers. The features on this dimension include 2nd person pronouns, conditional clauses, and prediction modals; these features commonly co-occur in certain kinds of student writing (e. g., If you wanted to watch TV a lot you would not get very much done).

Even in a particular study, the functional dimensions identified are either distinctive from others or share much in common with dimensions found in other MD studies. For example, among the four dimensions identified in Biber's (2006) study of university spoken and written registers, two of these are similar linguistically and functionally to dimensions found in earlier studies: Dimension 1 (Oral vs. Literate Discourse) and Dimension 3 (Narrative Orientation). In addition, the other two dimensions are specialized to the university discourse domain: Dimension 2 (Procedural vs. Content-focused Discourse) and Dimension 4 (Academic stance). In Dimension 2, the co-occurring 'procedural' features which include modals, causative verbs, 2nd person pronouns, etc. are especially common in classroom management talk and other institutional writing. The complementary 'content-focused' features which include rare nouns, rare adjectives, and simple occurrence verbs are typical of textbooks, particularly natural science textbooks. Dimension 4 which consists of features like stance adverbials (factual, attitudinal, likelihood) and stance nouns that-clause characterize teaching and management talk in the classroom. In Biber's (2004a) MD analysis of conversational text types, however, a

stance focused dimension is also identified. The difference between them is this stance focused dimension contains not only positive but negative loadings, giving it more diversified features. In this dimension, which is named 'stance-focused versus context-focused discourse', stance focused conversational texts were marked by the co-occurrence of that-deletions, mental verbs, factual verb - that-clause, likelihood verb - that-clause, likelihood adverbs, etc., while the context-focused texts had high frequencies of nouns and WH-questions, used to inquire about past events or future plans.

These issues raise a question for us to think about: whether or not the corporate blogs share any similarities or exhibit any differences to those underlying dimensions/text types of variation that have been identified, and what do those similarities and differences imply? Among those studies that have incorporated complete MD analyses of English registers, two parameters seem to be fundamentally important, regardless of the discourse domain: a dimension as associated with informational focus versus (inter) personal focus, and a dimension associated with narrative discourse. At the same time, these MD studies have uncovered dimensions particular to the communicative functions and priorities of each different domain of use. Although MD approach has been largely and successfully applied, its combination with SFL is still an innovation and somewhat tentative work. This theoretical framework will be applied in the current research, and what is presented in the next section focuses on the strengths and weaknesses of SFL and MD approach.

2.3.5 Strengths and weaknesses of SFL and MD approach

This section will focus on the strengths as well as some weaknesses of both SFL and Biber's MD approach to register variation. A review from this perspective aims to provide the rationale for the theoretical framework adopted for the present research project, i.e., the common grounds the MD approach and SFL share and a complementary role they play to each other support the combination of them and their adoption for the current research project. This section is started by looking at register firstly, and followed by a critical review of some controversial comments on both Biber's work and SFL theory, and finally an elaboration that the combination of the two may overcome the shortcomings.

As indicated in Section 2.2.1, register is typically described as the use-dependent variation of language (Gregory, 1967), or more specifically, “the linguistic features which are typically associated with a configuration of situational features” (Halliday & Hasan, 1976, p. 22). This ‘configuration of situational features’ is more explicitly elaborated both theoretically and practically. From a more technical and practical perspective, registers are said to be realized by a relatively high or low frequency of occurrence of particular lexico-grammatical features (Halliday et al., 1964; Matthiessen, 1993; Biber, 1988). Besides, the theoretical importance of linguistic co-occurrence has been emphasized by some researchers, such as Firth, Halliday, Brown and Fraser. Brown and Fraser (1979, pp. 38-39) observed that it can be “misleading to concentrate on specific, isolated linguistic markers without taking into account systematic variations which involve the co-occurrence of sets of markers”. Halliday (1988, p. 162) defines a register as “a cluster of associated features having a greater-than-random...tendency to co-occur”. All these claims, whether theoretically or practically oriented, suggest the compatibility of Biber’s MD approach and SFL. Apart from the correlation between the MD approach at the practical end and SFL at the theoretical end, it is also the location of text types (register) along the cline of instantiation which is the dimension immediately relevant to metafunctions (the interpersonal, the textual and the ideational; the ideational involving the experiential and the logical) that further justifies of this combination (Halliday & Matthiessen, 2004).

The feasibility of MD approach and SFL can be more fully revealed by looking into the advantages and disadvantages of both MD approach and SFL, thus providing us a more comprehensive picture of this combined framework as well as its rationality. For MD approach, it uses multivariate statistical techniques, e.g. factor analysis and cluster analysis, to analyse the co-occurrence relations among linguistic features, thereby identifying underlying dimensions of variation in a language. This approach is usually corpus-based, and requires the use of computational tools. The advantage of Biber’s MD approach over other methods used in register variation is that this method is more comprehensive than other approaches which have focused on a single register, a small number of linguistic features, or a smaller

number of texts (Biber, Conrad, Reppen, Byrd, & Helt, 2003b). In spite of this, MD approach has attracted some criticisms. In my point of view, some of them are reasonable, but some of them are less convincing.

Some weaknesses of Biber's approach have been pointed out by Ghadessy (2003):

> "since register is a cover term for any language variety defined in situational terms, a comprehensive analysis should tell more about how a text is created rather than focus on the co-occurrence patterns of certain language features; A comprehensive analysis of any register cannot "neglect the clause/sentence, which occupies an important position in any text, and should also consider other discoursal features"; by using MD approach, texts are analyzed as 'products' rather than 'meaning-making processes" (Ghadessy, 2003,P. 149).

I consider all these points are criticisms focusing on theoretical issues, rather than some constructive suggestions. It has been pointed out by Biber, Conrad, Reppen, Byrd, and Helt (2003a) that the strengths of MD analysis have not been acknowledged. I also doubt how a statistical method is able to reveal how a text is created, since the creation of a text is complex process and can hardly be quantified into some observed, correlated variables that the statistical method aims to describe. When it comes to the point that texts are analysed as 'products' rather than 'meaning-making processes by using an MD approach, I consider the analysis of texts as 'products' is simply a reflection of the research constructs and goals of MD analysis, rather than a flaw. I assume this critical point concerns different perspectives to view language: the choices as both process and as product that are interdependent on each other. In fact, what is behind any text is a complex selection of semantic options, and the product (text) is a representation of the process at a particular state. Any distinction between process and product may raise problems of its own. Andersen (2002, p. 200) describes the problem in this way: "the model describes wholes as emergent entities, generated by the activities of the parts, but does not specify the relationship between the new macro behaviour and the micro behaviours that created it." Although a text behaves like a unit, or like a whole, it is still very different from any of its parts. In all, it is a challenge to successfully account for the interaction of the parts in creating the language as a whole. And it is difficult to associate probabilities in an empirical approach with choices of language separated as either a product or a

meaning-making process. I consider the two views are inseparable: in the register analysis by using Biber's MD approach, texts are analysed as a product; from another angle, the different systems for meaning-making, or possible "channels" (e.g. speech, writing) could be thought of as being investigated through the distribution of a large set of linguistic features in them.

Besides, I also doubt how a comprehensive analysis of register can be realized in practice, when the term *comprehensive* has an "all-inclusive" meaning in it (Biber et al., 2003a, p. 151). The comprehensiveness of MD approach concerns the issue of the selection of language features for register analysis viewed from both an abstract and a practical perspective. Teich (2003, p. 26) has made a rather abstract comment on this issue:

> "when the common definition of register as part of the language system as a whole, realized by the relatively high (or low) frequency of co-occurrence of particular grammatical features, then what cannot be said with Biber's analysis is what the other grammatical options in the same systemic context are that might have been chosen, but were not. The only thing that can be stated is absence or presence of a feature, but not alternative lexico-grammatical renderings. Thus, a possibly useful complementary statement about which features have not been chosen that might have been potentially available, cannot be made. The choice of particular grammatical features can thus not be traced back to the language system".

In this abstract comment, it is not encouraged to pay great attention to the details of the selection of language features. Selecting or not selecting any lexical or grammatical features in an MD study may not hinder the analysis. This point coincides with those empirical support proposed by researchers specializing in factor analysis. Some detailed arguments concerning the optimal number of variables in a factor/dimension may support this:

- Six or seven indicators per factor and a rather small number of factors is considered as high over-determination of factors (MacCallum, Widaman, Zhang, & Hong, 1999)
- At least four measured variables for each common factor and perhaps as many as six (Fabrigar, Wegener, MacCallum, & Strahan, 1999, p. 282)
- A factor with fewer than three items might be weak and unstable (Costello & Osborne, 2011)

All these evidence suggests that results of a factor analysis/ an MD analysis would be more

accurate when each factor is represented by a moderate rather than a large amount of variables in the analysis. Therefore, in an empirical study, it is not true that the language features selected for an MD analysis should be as much as possible. It is not necessarily the case that the more comprehensive, the better.

Apart from the relationship between comprehensiveness of MD approach to the selection of language features, its comprehensiveness should also concern some underlying theoretical models for a language that the interpretation of results rely on, which is what MD approach lacks. Although this approach was initially developed for comprehensive research goals (e.g., Biber, 1988, 1995), and was said to provide the first and the most comprehensive investigation up-to-now to the study of linguistic variation in some registers, its lack in the support of a solid and systematic theoretical foundation is obvious. According to Teich (2003, p. 26), MD approach lacks an underlying model of language, as presented in the following ways:

> "(1) there is no explicit placement of the notions of functional dimension and communicative function in terms of linguistic stratum, (2) the distinction of form and function is not acknowledged with respect to the grammatical features, and (3) grammatical paradigms do not figure explicitly in the approach".

I agree to these points in that this critique raises our awareness of the importance of an underlying theory and model of language for a methodology. As an empirical approach, Biber's MD analysis in its own right does involve the interpretation of the results, and the interpretation of functionally important linguistic features was dependent on a large number of the previous research studies[6]. If relating the fundamental features of MD approach mentioned above to the notion of register as being part of the general linguistic system, it seems that a more fine-grained model explicitly revealing placement of system and text, function and form, different levels of abstraction in linguistic description (context, semantics, grammar), etc., can be integrated with MD approach. In order to identify linguistic features across text types more specifically as well as systematically, Halliday's SFL is an appropriate model to combine with MD approach, because it has met those requirements mentioned above. It has been

[6] These background studies include previous comparisons of spoken and written texts (see, e.g., the survey by Chafe and Tannen 1987), functional studies of particular linguistic features (e.g., Thompson 1983; Altenberg 1984), and descriptive grammars of English (especially Quirk et al 1985).

discussed in the previous sections that there are some common grounds shared by SFL and Biber's work, in terms of functional view of language, co-occurrence patterns of linguistic features and the corpus-based study in linguistic research. To be more accurate, Biber's MD analysis is rooted in Halliday's (e.g. 1964, 1978) SFL. The difference is that by using MD approach a register seems to be looked at from the text end as a set of texts, while in SFL, both text view and the system view on registerial variation are involved in a complex language system which play a complementary role and neither can do without the other.

When it comes to Halliday's SFL theory, it has been recognised as providing a very useful framework for viewing language as a strategic and meaning-making resource. The particular achievement of SFL theory was the demonstration concerning how the grammar of a language realized the choices with respect to each of the metafunctions, and how these in turn created the text with its register variables. More importantly, SFL theory "involves a theory about the nature of social life, a theory of language as a fundamental semiotic system involved in the shaping of social life, and a theory about the possibilities of social change" (Christie, 2004, p. 21). Such a theory, as what Halliday has always claimed, has much to offer to different social science disciplines, such as discourse analysis.

Ever since the foundation of SFL by Halliday, it has been developing evolutionarily rather than revolutionarily by Halliday and his followers, as well as getting some criticisms. The scope of SFL has been comprehensive from the beginning, and consistently developing from a comprehensive overview map of language in context towards a more detailed one (Matthiessen, 2007). This tendency has involved not only the focus moving from general to detailed issues, e.g., the description of the lexicogrammar of a given language, but "adding new semiotic dimensions to give a more multi-faceted view of language in context" (Matthiessen, 2007, p. 505). Although SFL approach has been widely accepted and used in holistic approaches in general, it has drawn some criticisms, many of which focus on three variables of context of situation: field, tenor and mode. For example, Hasan (2009, pp. 179-180) has critisized standard SFL applications of the terms 'field', 'tenor' and 'mode' as vague, lacking "checkable" criteria, and relying on "common sense". System networks for

contextualization remain “at a nascent stage”, although they “actually systemize the realization-activating contextual features and attempt to relate context to wording via meaning, which acts as the interface between the two” (Hasan, 2009, pp. 181-182). Besides, some more criticisms on the three variables have also been summarized by Van Dijk (2008, p. 54):

1) “ Halliday, and then later other SFL linguists, borrow a notion of context from Gregory and other UK linguists that is vague, heterogeneous, terminologically idiosyncratic and theoretically confused, namely the triple of ‘field’, ‘tenor’ and ‘mode’. With small changes, this conception has barely changed in nearly forty years, although it produces numerous problems for the theory of the relations between text and context”.
2) “Because of the rather arbitrary nature of these three “variables” defining contexts, the mapping of such contexts on (three) functions of language (ideational, interpersonal and textual), and on the language structures controlled by them, also remains arbitrary, incomplete and confused. This not only shows in the theory but also in the analyses of language use in SFL”.

It should be admitted that these criticisms have addressed some aspects of the problems in SFL; however, these criticisms are also limited in some way. In general, all these comments are concerned with the modeling of the concept of context of situation within SFL. The nature of these three ‘variables’ defining contexts is labelled as being vague, confusing and incomplete; however, I assume Halliday perceives these as being natural. Moreover, the vagueness and incompleteness of three variables could be thought of as a limitation rather than a flaw of SFL. Within SFL, there have always existed some interests on modeling context as a system network, in which the relevant parameters of three variables or the patterns of the stratification of a register, etc… are identified. For example, there are discussions on features from different parameters of context in various sources in the SFL literature (Butt, 2004; Cloran, 1987; Hasan, 1999, 2009; Martin, 1992), and there are also some recent work on context inspired by Hasan’s network for ‘field’, e.g., Butt’s framework on ‘Parameters of Context’ (e.g., Butt, 2004; Hasan, 1999, 2009). There is so far no unified descriptive framework of the features of context of situation. In spite of this, it may not be a problem for the current research,

because this research is based on certain component parts of the primary systems and features for Field, Tenor and Mode.

In terms of the interrelations of three variables, their interdependence are said to be "partial"; in specific, "what happens typically is that they display a mutual prehension; the echoes of a choice in one are found to some extent in the choice of the others" (Hasan, 1999, p. 245). The close tie of three parameters is an attribute in nature, rather than a vague and confusing phenomenon. They are closely related aspects of context consisting of such elements like topics, people involved and the medium of communication. To some extent, all these three aspects do not constitute a complete picture of a register, but present a relatively broad category of register, in comparison with some other researchers, such as Hyme[7] (1974) who proposed a complex model consisting of many categories. Halliday (1978) sees mode as covering Hyme's channel, key and genre, i.e., the message being passed through speech, writing, signing, etc (channel); the evaluation of message whether the message is good or bad, sorry or impressive (key); the message form intended may be a narration, sermon, etc (genre).

Some other criticisms are from Hasan (1999, p. 225) who highlighted some shortcomings of the SFL model in terms of contextual/registral constancy. They are briefly listed below:

1) "the prevalent contextual descriptions in SFL are based on an assumption of contextual constancy across a given text"…;
2) "While the claim of contextual/registral constancy is empirically validated in the majority of cases, it is not universally true. It is possible to find cases where the integrity of a text is able to survive certain kinds of contextual/registral changes";
3) "This being the case, the question arises: what would one mean by "coherent with respect to the context of situation" in the latter type of cases? Current SFL models possess no satisfactory means of answering this challenge or of specifying the nature and character of those contextual and/ or registral changes which do not disturb the

[7] Hymes' (1974) taxonomy comprises the categories Setting/Scene, Participants, Ends, Act sequence, Key, Instrumentalities, Norms, and Genres, which together form the acronym SPEAKING. This model has been widely applied to characterize novel or exotic speech communities, serving as a preliminary descriptive framework that draws the researcher's attention to aspects of the speech situation that may assist in interpreting linguistic phenomena of interest.

> unity of the text, nor can they specify where i.e., in what kind(s) of social situation, such changes are most at risk".

These criticisms raise a question concerning how and to what extent those assumptions proposed in a theoretical framework can be proved right. I suppose empirical studies can be used to verify theories, and vice versa, their intimate relationship guarantees that theory reflects underlying realities. The theory should be rooted in concrete data, or evidence, which is able to "provide a compass that can keep the theory generation process on course" (Alvesson & Karreman, 2011, p. 2). For example, Biber's claims on form-function relationships may work for the benefit of theoretical reasoning in SFL, in that the two points listed below illustrate a variety of contextual changes, which contribute to a more reasonable judgment of the shortcomings mentioned above. Firstly, in many cases, there are multiple functions associated with the collection of linguistic features on a dimension, rather than a single communicative function underlying each grouping of co-occurring linguistic features. Secondly, it is not the real situation that there is a single communicative function associated with each linguistic feature. Rather, in many cases, features can have somewhat differing functions in different kinds of text; further, features can have functions defined at different levels of generality. Overall, no matter how dynamic the relations between empirical studies and theoretical frameworks are, their mutual interactions and influences cannot be neglected. The issue of the combination of SFL and Biber's MD approach has been discussed by some researchers (Gardner, 2008; Nesi & Gardner, 2012; Nini & Grant, 2013) who emphasize not only the importance of SFL theory, but the reliability of Biber's method. For example, Nini and Grant (2013) comment that Biber's method is a reliable way of computationally calculating registerial variation, as claimed in SFL. And Gardner (2008) argues that without the contextual work on text types or SFL analysis into genre families, it would be difficult to make sense of the results of MD analysis. Although Halliday and Biber may differ significantly in their methodology, both of them are concerned with the identification of lexico-grammatical features that occur in specific social contexts. In the current research, the combination of SFL and Biber's MD approach to register variation is adopted as the theoretical framework providing the rationale for conducting my research.

2.4 Blogs and corporate blogs

2.4.1 Background

As mentioned in the introductory chapter, the use of blogs (the shortened version of Weblog or Web log) is relatively a new phenomenon in the use of Internet and computer mediated communication (CMC). From the first launch of a blog between 1994 and 1998 (Gurak et al., 2004) to its move into the mainstream in recent years, some changes have been taking place. At first, the blog was used for those personal bloggers to publish online journals (Quible, 2005). Later on, blogging was gradually becoming a commonly-used tool in different disciplines, such as politics, psychology, etc., thus giving rise to the classification of different blog sub-types, including corporate blogs (Cowen, 2004). In the blog family tree, the main distinction of blog types has been made on the basis of content types by different researchers from non-linguistic areas. Krishnamurthy (2002) classified blogs into four basic types according to two dimensions: personal vs. topical, and individual vs. community. The topical blog was further divided into two categories by Herring et al. (2004): filter blogs, which contain a blogger's comments on articles or postings from other web sites that are linked to the blog itself; and klogs, short for knowledge-logs, which are basically informational websites on a particular topic written by an expert. The klogs proposed by Herring et al. (2004) is also thought of as a replacement of the blog type named notebook, one of the three basic blog types by Blood (2002). According to Blood (2002), blogs include filters, personal journals, and notebooks. Among these three types, the content of filters is external to the blogger (world events, online happenings, etc), and the content of personal journals is internal (the blogger's thoughts and internal working), while notebooks are usually long focused essays. Bar-Ilan (2004) listed three types of blogs based on content: associative, personal and self-expressive, and topic oriented. In his categorization, topic blogs refer to blogs which aim at talking about topics relating to a hobby or to the author's profession or business. These categorizations of blogs are based on their content types, which mainly correspond to a participant's taxonomy and the shared knowledge deriving from different participants or social categories that are pertinent to the communication. It is assumed that these categorizations constitute the central attributes of blogs, and they are in

parallel with another set of variation, i.e., linguistic variation which is conditioned by situational context. In the current research, it is the functional linguistic variation in corporate blogs that will be explored.

During the time when the blog type grows increasingly abundant, the definitions of blogs also evolve gradually. Most definitions are technologically oriented. A vast majority of researchers commonly accept the definition which describes blogs as frequently modified web pages in which dated entries are listed in reverse chronological sequence (e.g., Bortree, 2005; Buckingham & Willett, 2013; Kelleher & Miller, 2006; Kwasnik et al., 2005; Schmidt, 2007). Besides, it has been claimed in some thematic-based definitions that blogs include posts similar to those found in a diary or a personal journal. Some specific examples are presented below:

- "an on-line journal where the author keeps a running account of whatever she's thinking about" (Suitt, 2003, p. 30);
- "a web diary" (White, 2003, p. 10);
- "virtual diaries created by individuals and stored on the web for anyone to access" (Sharda & Ponnada, 2008, p. 157);

Many years after the first launch of a blog between 1994 and 1998, many scholars have characterized blog as a genre. Among them, Miller and Shepherd (2004) are the first to describe the blog "as a genre that addresses a timeless rhetorical exigence in ways that are specific to its time" (Miller & Shepherd, 2004, p. 17). They consider journal and the diary, along with the newer electronic genres of the home page and the webcam to be the branch of the blog family tree. Blood (2002) has commented that the blog represents a new digital phenomenon without any antecedents, which seems to imply it is a new emerging genre. Herring et al. (2004, p. 9) have argued against some of those points by Blood (2002) by saying that "Blood's claim about the origins of the blog which is based on the assumption that blogs are link-centered Web content misrepresents most blogs at the present time". In fact, as a very popular blog type, personal journal blogs contain few links and do not focus on web content. In addition, Herring and her colleagues have also sensed the hybridity of blogs, and explored the

categorization of blogs, which have shed light on its place in the genre ecology of the Internet. To some extent, important contribution to research into blogging genre is from them (e.g., Herring & Paolillo, 2006; Kwasnik et al., 2005), whose study is a piece of pioneering work to the systematic description of blogs. Following Herring's work, some more linguistic-oriented studies on blogs and corporate blogs have been conducted, including those from which the gap to fill in the current research can be indicated. They will be reviewed in the following section, with the definitions to blogs and corporate blogs presented briefly at first.

In terms of corporate blogs, the definition appears relatively lately. The corporate blog has been defined as "a blog published by or with the support of an organization to reach that organization's goals" (Fredrick, 2004), or "an informal, easily maintained way to regularly communicate with customers and employees" (Heathfield, 2010). This new media is not like traditional business media (e.g. corporate websites) that require lots of time for preparation; rather, it involves frequently updated articles of opinions, news and information. There are basically two types of corporate blogs – internal and external – depending on the intended audience and purposes (Stuart, 2006). Internal corporate blogs are used to facilitate communications within a corporation by acting as a content management system, while external corporate blogs, as the name suggests, are used for communication between the corporation and the public, which includes the media, customers, and business partners.

2.4.2 Linguistic studies on blogs and corporate blogs

As a recently emerging language phenomenon, blogs have not received much scholarly attention. Yet, some constructive studies have been conducted on blogs, and most of them have been taken from different perspectives, focusing on various topics, and mainly taking a non-linguistic approach. Some researchers have presented empirical findings on the socio-demographics of bloggers, as well as their motivations and habits (e.g., Huffaker & Calvert, 2005; Kwasnik et al., 2005; Lenhart & Fox, 2006; Nowson et al., 2005), which indicate the dominance of the personal journal-type blog. In addition, extensive research studies have been conducted on particular practices, e.g. blogging about political topics (e.g., Bahnisch, 2006; Park, 2009; Resnick, 2011; Singer, 2005; Wallsten, 2007), blogs and journalism (e.g.,

Haas, 2005; Lasica, 2001; Robinson, 2006; Vobič, 2007), blogs for education and e-learning (e.g, Brownstein & Klein, 2006; Divitini & Morken, 2005; Williams & Jacobs, 2004), blogs for organization communications, i.e., corporate blogs (e.g., Charman, 2006; Kelleher & Miller, 2006; Lixia, 2004; Yang, Kang, & Johnson, 2010). In comparison with this, very fewer studies have explored the linguistic properties of blogs and corporate blogs. However, this does not mean the universe of blog texts involves similar situational characteristics of blog genres. Rather, the complexity of blog language shall be viewed from their dramatic differences in terms of their linguistic characteristics.

The complexity of blog language has been witnessed and explored by some researchers from a linguistic perspective. Among these studies, the central issue is the status of blog discourse in relation to other genres. The blog has been considered as a hybrid genre in nature, the language of which has much in common with spoken and written language (Aleknavičiūtė, 2009; Nilsson, 2003; Tavosanis, 2006) or with personal homepages and asynchronous discussion forums (Herring et al., 2004). In addition, it was also found that there are some differences of the formal/informal language use in blogs, email and school essays (Nowson et al., 2005), and a marked variation between traditional academic genres and academic weblogs in the use of first- and second-person pronouns (Stuart, 2006). All these studies show that blog language is characterized by hybridity articulated in many different ways.

In terms of the status of blogs in relation to both spoken and written language, Nilsson (2003) made a useful starting point on the study of blog language in his paper named “The Function of Language to Facilitate and Maintain Social Networks in Research”. In this paper, Nilsson applied Crystal’s (2001) list of differences between written and spoken language to blogs: Boundedness and Dynamicity; Synchronicity; Paralinguistic Cues; Constructions; Communicative Functions; Ability to be Revised; Unique Communicative Features. The result of the study shows that the language of blogs has much in common with both and fall somewhere in between, instead of falling completely under either medium.

After Nillson’s pioneering research in 2003, the position of blog language between spoken and written language was explored more broadly in different sub-types of blogs, and in the

languages other than English. Tavosanis (2006, 2007) has conducted a series of studies on the language in Italian Blogs, a preliminary result of which is the language of blogs is not restricted to the more informal levels of expression. Instead, blogs may include many kinds of written language, from simple personal notes to literary prose or poetry. Aleknavičiūtė (2009) uses genre analysis methods to describe and compare the linguistic features of political blogs in English and Lithuanian. What is explored in this study is more than the position of blog language between spoken and written language; it concentrates on a wide range of topics, from the analysis of linguistic features, to its formal representation and cognitive structure, which are discussed in relation to the contextual factors influencing the weblog. The results of this study include: blogs by British and Lithuanian Members of the European Parliament (subsequently MEPs) are not same linguistically; blogs by British MEPs combine both spoken and written language; some written language features include very long complicated sentences, and the formal vocabulary (i.e. the use of a field specific vocabulary, usage of impersonal structures or passive voice rather than the first person singular). The spoken language features include short expressive interrogatory or exclamatory sentences, and the informal vocabulary, such as phrasal verbs or colloquial words. In comparison with blogs by British MEPs, blogs by Lithuanian MEPs are more dominated by the informal vocabulary, and show a more vivid, emotional, and a more distinct personal stylistic variation.

Apart from being positioned between spoken and written languages, blogs have been viewed from a more complex and systematic perspective, particularly after Herring's (2004) pioneering work which comprehensively describes the characteristics of blogs. The study by Herring et al. (2004) has made an important contribution to the research into blog genre. The results of their genre analysis revealed not only the linguistic characteristics of blogs, but also the whole picture of blogs (e.g. blog author characteristics, distribution of blog types according to their primary purpose, structural characteristics of the home pages of the blogs). In relation to blog language, the results of the study suggest that blogs have been identified as containing both elements of previous Web genres, such as personal home pages, and elements from previous offline genres, such as diaries, editorials and newsletters, thus forming a hybrid and evolving genre. After this pioneering work, a genre perspective has been incorporated into many

subsequent studies. Or even in some cases, a genre analysis approach was used for the investigation of blog language in relation to other genres, as well as the social and cultural factors influencing the blog language. In particular, some subsequent linguistic studies have mainly focused on blogs in some more particular domains, such as academic blogs (Stuart, 2006) and corporate blogs (Puschmann, 2010a, 2010b), or by using the multi-dimensional analysis approach (Grieve et al., 2011). There is also a study which presents a descriptive and comparative research on cohesion in blogs (Hoffmann, 2012). An overview of the research on blogs is displayed by Myers (2010), who identified specific discourse characteristics and demonstrated the wide range of perspectives from which blogs can be analysed.

Instead of focusing on non-cohesive lexical and grammatical language features, Hoffmann (2012) presents a corpus-based descriptive and comparative research on cohesion. This study employs a self-compiled corpus of personal blogs, which is named the Augsburg Weblog Corpus (AWC), and examines how the use of cohesive devices varies in blog entries and comments. It also investigates the degree to which blogs resemble prototypical written monologues or spoken dialogues. The results of this study show that the distribution of grammatical and lexical cohesion in blog entries and comments is similar, and the cohesive profiles in blogs as a whole are in many respects similar to those in written monologues rather than those in the spoken dialogue. More specifically, blog entries and comments have been shown to feature narrative and argumentative text types which can equally be found in written monologues (e.g. fiction and scientific articles) as well. Overall, both of cohesive and discursive interactions in blogs are extremely constrained.

There are also some research studies on specific types of blogs. In Stuart (2006), a genre-based linguistic study of academic weblogs was conducted, the purpose of which is to reveal how much variation there might be between traditional academic genres and academic weblogs. Based on a corpus consisting of 39 academic weblogs with 16 million words, the linguistic features in this corpus were compared with the ones from Biber's (1988, 1995) earlier study of academic prose. The result of this comparative study shows that there is one area in which the two genres have marked variation, i.e., the use of the first and the second person

pronouns. This reflects the diary/journal aspects of weblogs and the fact that weblog authors are very aware of their audience. Stuart (2006, p. 391–392) also analysed the communicative purposes of 496 blogs within academic organizations, and came up with a list of 19 categories which could be further classified into four more general categories (Table 2.4). These general categories can be considered to form “a subset of genres within the larger system of academic and educational genres which are inter-textually and inter-discursively linked within the boundaries of academic and educational settings”.

Table 2.4: Uses of blogs in academic institutions (Stuart, 2006, p. 392)

Author	Student	Teacher
Researching	PhD/masters Weblogs	Academic announcements, critique, debate
Learning/teaching	Student weblog as coursework Student group discussion tool	Class weblog Publication of texts, links or commentary to seed discussion

Apart from academic blogs, corporate blogs have also been explored from a linguistic perspective. In Puschmann’s (2010a) thesis named ‘the corporate blog as an emerging genre of computer-mediated communication: features, constraints, discourse situation’, a broad range of issues have been explored. In Puschmann (2010a), a basic stand point is “the wider socio-pragmatic context of blogging must be examined before any narrower observations on corporate blogging or on micro-linguistic phenomena can be made” (Myers, 2010, p. 25). He first looked into the technical and pragmatic aspects of blog, then the corporate blogging as an emerging genre, and at last the micro-linguistic aspects of corporate blogs. In general, Puschmann’s (2010a) work has provided us a very wide vision of the larger context of blogs, or larger extra linguistic factors on linguistic dimensions in his study. And these aspects are explored on the basis of a multitude of disciplines, such as sociology, psychology, cultural and media studies, computer sciences, literary studies that may influence or help construct a theoretical framework, and by using a diverse range of approaches and methodologies. As for the micro-linguistic aspects of corporate blogs, a comparison of corporate blog language with Business English and Standard English is made and the comparison is restricted entirely to

lexical aspects, i.e., the use of *nouns* and *Pronouns.* The result shows that corporate blogs share lexical similarities both with the characteristics of Business English and Blogs. In another paper by Puschmann (2010b), the function and use of first person and second person pronouns (inter-personal pronouns or IPPs) in blogs maintained by several major U.S. companies for marketing and public relations purposes was examined. Several strategies of IPP use in corporate blogs were identified: nobody (The blogger avoids self-reference and does not address an audience); only you (The blogger avoids self-reference and addresses an audience); only me (The blogger refers to himself or herself and does not explicitly address an audience); only us (The blogger refers to the company, the blog team etc. and does not explicitly address an audience); you and I (The blogger refers to himself or herself and addresses an audience of 1-to-X people); you and us (The blogger refers to the company, the blog team etc. and addresses an audience of 1-to-X people).

Overall, the linguistic studies on blogs and corporate blogs presented above provide support for the view on blogs/corporate blogs as an emerging genre, as well as some important information about their linguistic features and functions; however, our understanding of the linguistic variation in corporate blogs remains limited in particular ways. Previous studies on corporate blogs focus on individual linguistic features; therefore, it is difficult to know how features interact and what the possible relations between forms (linguistic features) and functions are. In Puschmann (2010a), the results of the quantitative analysis is weakly linked to the communicative qualities of corporate blogs, although it was claimed that "this quantitative data supports the broader claims made about the communicative qualities of corporate blogs" (Puschmann, 2010a, p. 109). On the one hand, Puschmann's work has provided a very wide vision of the larger context of blog and CB communicative function, such as the formal, technical, and pragmatic aspects of blogging, issues of corporate communications on the Internet, origins and perceived advantages of corporate blogging, a typology of corporate blog subgenres, pragmatic aspects of corporate blogs, etc. This wide vision is characterized by the focus on a wide range of extra linguistic factors, and the construct of a theoretical framework consisting of a multitude of disciplines, such as sociology, psychology, cultural and media studies, computer sciences, literary studies. On the other hand,

the quantitative analysis in Puschmann (2010a) looks into frequent nouns in CBC (Corporate Blogging Corpus), and Pronoun frequencies in CBC, BNC (the British National Corpus) and BEC (the Business English Corpus). Therefore, the application of quantitative methods for the comparison of noun and pronoun use is weakly linked to the wide range of extra-linguistic factors.

Given these limitations, it would be necessary to have a more systematic perspective on linguistic variation in corporate blogs. Such a perspective will emphasize the investigation of a large number of linguistic features while still making functional interpretations of language use. An MD approach, which involves the use of multivariate statistical techniques to study register variation in language corpus based on the grouping/co-occurrence patterns of linguistic features is directly applicable to understanding variation in corporate blogs. This approach has played an important role in shifting researchers' attention from an individual analysis of particular linguistic features to viewing language variation as a continuum. In the present research project, an MD approach is firstly employed to identify the principal dimensions/factors and text types/clusters of corporate blogs, and then the salient linguistic differences across different text types are identified in the SFL model.

2.5 Summary

In this chapter, I have reviewed the concept of 'register' from two perspectives: theory and practice. From a theoretical perspective, I provide an overview of the theoretical origins of register variation, and then review register in SFL, which concerns dimensions of stratification, metafunction and instantiation in general, and three metafunctions (ideational metafunction, interpersonal metafunction, textual metafunction) in particular. Besides, the language features at the lexicogrammar level, and the semantic analogue of the variables of the social context (field, tenor, mode) that are closely related to three metafunctions are reviewed. From a practical perspective, Biber's MD analysis framework for the investigation of register variation, as well as some previous MD studies is discussed. Following these, a discussion of the strengths and weaknesses of SFL and MD approach is provided, which takes a more objective and critical perspective on these components in the theoretical framework, and thus, provides

a better foundation for their integration and combination in the theoretical framework. Finally, a review of previous studies on blogs and corporate blogs from a linguistic perspective is provided. All the background information from these earlier studies helps us highlight the problem and the gap in the literature, from which research hypotheses and research questions to investigate or guide the current study can be framed. Before the introduction to the methodological issues (Chapter 4), and the presentation/discussion of the results of the study (Chapter 5 & Chapter 6), a review focusing on the lexical and grammatical features selected for the MD analysis is presented in the following chapter.

Chapter 3: Linguistic features selected in the multi-dimensional analysis

3.1 Introduction

The selection of linguistic features plays an important part in the multi-dimensional analysis. This chapter provides a detailed description of the linguistic features selected for the MD analysis. Based on a review of previous theoretical and empirical work on register variation and related lines of inquiry such as the study of the differences of written and spoken communication, it is decided that 32 linguistic features could be selected for the current research. This is based on a proved assumption that they could be significant statistically in the current analysis. The linguistic features selected for this MD analysis are from a wide range of 11 grammatical/functional categories (Biber, 1995, pp. 94-96): (A) Tense and aspect markers (1-3), (B) Place and time adverbials (4-5), (C) Pronouns and pro-verbs (6-11), (D) Questions (12), (E) Nominal forms (13), (F) Passives (14-15), (G) Stative forms (16), (H) Subordination features (17-24), (I) Prepositional phrases, adjectives, and adverbs (25-27), (J) Lexical specificity (28-29), (K) Modals (30-32). What follows is a detailed description of these linguistic features.

3.2 Features selected in the multi-dimensional analysis

The selection of linguistic features for the multi-dimensional analysis in this thesis is primarily based on the functional studies by Biber (e.g, 1988, 1995), in which a subset of features define the dimensions identified from his multi-dimensional model of register variation. The identification of these functionally important linguistic features from Biber (e.g., 1988) was largely dependent on the body of previous research studies, which include “previous comparisons of spoken and written texts, e.g., the survey by Chafe and Tannen (1987), functional studies of particular linguistic features (e.g., Altenberg, 1984; Thompson, 1983), and descriptive grammars of English (especially Quirk et al., 1985)” (in Biber, 1995, p. 96). Among these different groups of relevant literature that the MD analysis relies on, I will particularly

focus on the work by Chafe (1982), Biber (e.g., 1988, 1995), and Quirk et al. (1985), although there could be a wide range of overlapping between them. On the basis of the survey of previous studies, 32 linguistic features have been selected from Biber's (1988) initial 67 linguistic features, which have also been grouped into classes based on their grammatical functions. The 32 linguistic features from 11 classes of grammatical functions include: 1. Past tense; 2. Perfect aspect; 3. Present tense; 4. Place adverbials (e.g., above, beside, outdoors); 5. Time adverbials (e.g., early, instantly, soon); 6. First-person pronouns; 7. Second-person pronouns; 8. Third-person personal pronouns (excluding it); 9 Pronoun it; 10. Demonstrative pronouns (that, this, these, those as pronouns); 11 Indefinite pronouns (e.g., anybody, nothing, someone); 12 Direct WH questions; 13 Total nouns; 14. Agentless passives; 15. By-passives; 16. Existential there; 17. Infinitves; 18. THAT relative clauses controlled by a verb; 19. THAT relative clauses controlled by an adjective; 20. THAT relative clauses controlled by a verb; 21. WH relative clauses; 22. Concessive adverbial subordinators (although, though); 23. Causative adverbial subordinators (because); 24. Conditional adverbial subordinators (if, unless); 25. Total prepositional phrases; 26. Total adjectives; 27. Total adverbs; 28. Type-token ratio; 29. Mean word length; 30. Possibility modals (can, may, might, could); 31. Necessity modals (ought, should, must); 32. Predictive modals (will, would, shall). What follows is a brief review of each linguistic feature[8], which is expected to follow the principle that "features can have somewhat differing functions in different kinds of text, and functions defined at different levels of generality" (Biber, 1995, p. 135).

3.2.1 Tense and aspect markers (1-3)

In this section, three linguistic features selected from the category of tense and aspect markers are introduced: 1) past tense, 2) perfect aspect, and 3) present tense. The grammatical functions of these linguistic features concern "aspect", which reflects "the way in which the verb action is regarded or experienced with respect to time", and "tense" that is relative to the time of utterance as well (Quirk et al., 1985, p. 188). What distinguishes the two categories are the types of constructions realizing the actions, i.e., the two types of aspect constructions

8 The sample sentences illustrating each of the 32 linguistic features are either from my corpus or Quirk et al. (1985).

including the “perfect” and the ‘progressive’, and two types of tenses of the verb including “present” and “past” (Quirk et al., 1985, p. 189).

As for the discourse functions of the three linguistic features, there are some variations and overlapping among them. In particular, both past tense verbs and perfect aspect verb forms are features marking “narrative” discourse: past tense verbs have been considered as a primary surface marker of “narrative” discourse (Biber, 1988, p. 223), and the co-occurrence of past tense verbs, and third person pronouns marks the narrative and descriptive discourse, as well as certain kinds of academic writings (Biber, 1986). In comparison with the two linguistic features mentioned above, ‘present tense’ has distinctive discourse functions associated with the personal involvement. “Involvement” is a term that Chafe (1982, p. 45) has used to refer to the expression of the speaker's “richness of his or her thoughts” or “experiential richness”. This term has been more often used by some linguists (e.g., Biber, 1988; Chafe, 1982; Tannen, 1982a) to describe the differences between spoken and written languages. For example, Chafe (1982, pp. 38-43) offers a group of features to explain the differences between spoken and written languages. Spoken language is more fragmented and involved, and therefore, many loosely joined clauses and many involved features are used. In contrast, by using many participles, attributive adjectives, nominalizations, and passives, academic texts are more integrated. According to Biber (1988, p. 105), texts characterized by personal involvement (Dimension 1) are “verbal, interactional, affective, fragmented, reduced in form, and generalized in content”. And ‘present tense’ is among bundles of features taken from Dimension 1 in Biber (1988).

Apart from the discourse functions of the three linguistic features, their grammatical descriptions are also reviewed. As suggested by the term, past tense verbs are used with reference to past events, as illustrated in sample sentences 3.1, 3.2 and 3.3. Besides, they can also be used to refer to present and future time (Quirk et al., 1985, pp. 183-188). In particular, they can be used as an alternative to present tense forms in indirect speech or indirect thought, when the past tense verb in the reporting clause “inflects” the tense of the verb in the subordinate clause as in *Did you say you have/had no money?* They can also be

used as a more polite variant of the present tense in the sentence like *I wonder/wondered if you could help us*, the so called “attitudinal past” (Quirk et al., 1985, p. 188). Finally, the past tense is used in certain hypothetical clauses, especially if-clause, to express a “hypothetical” thought, as shown in sample sentences 3.4 and 3.5.

3.1) My parents *started* out with very little in their pockets and *built* a restaurant business. I *started* out working with one hotel and now I have been blessed with many more, almost 3800. I *told* my daughter, Debbie, "I'm blown away." I never *imagined* all this. (file Marriott20120418)

3.2) LinkedIn employees *submitted* a wide range of ideas from internal productivity tools to LinkedIn site features to using LinkedIn for charitable opportunities.(file LinkedIn20120829)

3.3) Before Alexander Graham Bell *invented* the telephone, (June 2, 1875) public safety *was* served by town criers. A town crier would walk the streets of a town and cry out for help in emergency situations. In the 1950's, independent telephone companies *were* very common in the United States. If you *wanted* the police, you *dialed* the police station. If you *had* a fire, you *called* the fire department. If you *needed* any emergency help, you *dialed* the individual you needed, or you *could* dial “0” and get the operator. Then he or she would ring the persons you were calling for. (file AVAYA20120216)

3.4) While Randy acknowledges that tiered storage allows for greater consolidation by managing the variations in storage requirements, he does not believe that the "single box for everything" concept is practical. I would agree with Randy on all these points if you *were* not using storage virtualization. (file HITACHI20120518)

3.5) This is just a quick snapshot of the overall picture. I wish I *could* single out everyone who deserves it. But I'd love if you *could* share any stories about exceptional associates you've met in our hotels. (file Marriott20120509)

As opposed to the simple past which usually indicates that the action is finished, the present perfect indicates that the action has continued up to the present time and may even continue into the future. The two verb tenses are not interchangeable; yet, the present perfective relates the action more directly to the present time (Quirk et al., 1985). In addition, the past perfect has the meaning of “past in the past”, and can be regarded as preceding the present perfective or the simple past (Quirk et al., 1985, p. 195), as illustrated in the following sample sentences:

3.6) Looking at the study results a bit closer, 42% of mobile users *have clicked* through on mobile web ads and 37% *have clicked* through on ads in mobile apps. Furthermore, men are more likely to click on in-app ads than women (42% vs. 32%).Keep in mind, the study reports users who *have clicked* on mobile ads at some point in time. That doesn’t mean that they click on mobile ads frequently or if *they’ve even clicked* on a mobile ad more than once. Perhaps more important is the finding that the majority of mobile users *have never clicked* through on a mobile ad at all. (file Corporateeye56)

3.7) It was such an amazing feeling to know that my hard work was finally paying off. I *had* already taken the job at Spirit when I was told I *had won* the scholarship from Delta. Even though I was working for a competitor, Delta was true to its word and worked with me and my training/work schedule at Spirit. (file Delta20120323)

Quirk et al. (1985, p. 179-181) make a distinction between three different meanings of the simple present tense: the state present, the habitual present, and the instantaneous present. Closely related to the three different meanings of the simple present tense is a principle proposed by Greenbaum and Quirk (1990), which concerns the distinction between stative and dynamic aspect of English verbs: stative verbs include the primary verbs like “be, have, think, want, live”, and the verbs of perception and bodily sensation, such as “see, hear, taste, and hurt”. Leech (1971, p. 4) has defined “a state” as being lacking in defined limits, as illustrated in sample sentences 3.8 and 3.9. Besides, what has defined as event verbs, activity verbs and process verbs by Leech (1971) could be categorized into dynamic verbs (Greenbaum & Quirk, 1990), which are usually confined to habitual/event/instantaneous present and past, as illustrated in sample sentences 3.10 and 3.11.

3.8) Military service *is* something near and dear to me and my Delta colleagues. Many of us have served or *have* family members that *are* in active service or who have served previously. I’ve found that this *is* the case with our customers too. The Honor Guard conducts plane side flag ceremonies with specially-themed ground equipment as these fallen heroes *are* moved onto Delta aircraft. (file Delta20120723)

3.9) But I *believe* it’s the domain of the ignorant to float the dismissive and tired chestnut, “Aww, ignore all that crap, it's been said before. Just go climbing.” Not because you shouldn’t go climbing (you should), but because the sentiment fails to recognize the inherent importance of dialogue. It *makes* us *think*. Based on the responses to Beal’s post, it touched some nerves and got people talking - and thinking. These are good things. (file patagonia20120420)

3.10) I found that being with Steph and Arnaud is a bit like how I imagine yoga camp being, but in a good way. They don't drink coffee or eat red meat, they blanch their own almonds and they definitely don’t pack crag beers. Steph writes poetry and Arnaud reads psychology magazines. So I was on my best behavior, drinking tea, eating fruit and staying hydrated with water instead of pilsner. I read 1.5 books and only checked my email once a day. (file patagonia20120418)

3.11) From now until March 18th, you can *go* to Whole Planet Foundation’s Facebook page for a chance to win a trip for two to India. (You’ll need to have your own Facebook account and “follow us” to enter.) You’ll be asked to *write* a short essay, no more than 350 words, *describing* how a small change you have made has had a big impact on the world around you. Feel free to include links to web pages, Facebook pages or other online locations that help tell the story. (file WholeFoods20120222)

Additionally, three uses of simple present are also mentioned by Quirk et al. (1985, pp. 179-181): simple present referring to the past, simple present in fictional narrative, and simple present referring to the future. Simple present referring to the past is used in narrative style, characterized by describing the past as if it is happening now, as shown in the sample sentence 3.12. Simple present in fictional narrative differs from historic present, in that events narrated by means of the historic present are real, and those narrated by the fictional "historic present" are imaginary. Furthermore, in simple present referring to the future, the event is expected to happen in the future, though occurring with time-position adverbials in the present. Usually the future use of the simple present is much more common in dependent clauses, particularly in conditional and temporal clauses, as shown in sample sentences 3.13 and 3.14.

3.12) Although central government is not generally thought of as being very innovative, I was pleasantly surprised to *hear* that a number of departments *were starting to embrace* newer ways of communicating with the general public, including greater engagement with social media channels. (file rs20120525-2)

3.13) But I do it so I know and I do it so that I always *know* exactly what *will happen* if and when I recommend something like that to my clients. (file chr20120612)

3.14) We've all got to do our bit to save Greece he declared and I've worked it out, if we all *drink* 89,000 of these tonight and tomorrow night, … then we'll get Greece back on its feet! (file rs20120330)

3.2.2 Place and time adverbials (4-5)

The previous section shows that the tenses could be presented in a variety of ways; however, they could not simply work without the help of time adverbials; rather, they need the "cooperation of time/place adverbials or context" (Declerck, 1991, p. 254). In this section, two linguistic features, i.e. 4) place adverbial and 5) time adverbial, as well as their semantic roles are focused on and introduced.

Both of the two features have been found in Dimension 3 (explicit vs. situation-dependent reference) in Biber (1995), a study which develops further Chafe's (1982) finding of linguistic features and their functions, as well as building the factorial structure of multi-dimensional paradigm. In Dimension 3 (Biber, 1995), time adverbials and place adverbials are used for locative and temporal reference in the actual physical context of the discourse. In Quirk et al.

(1985, pp. 479-482), several semantic relations expressed by adverbials in relation to physical space/place and time have been distinguished. In specific, the semantic roles of place adverbials include position, which is normally associated with verbs referring to stasis, as well as motion, direction, which may refer to directional path with or without locational specification, and goal which involves a positional aspect of direction.

In the literature, it has been widely assumed that the passing of time is conceived of in the same way as movement through space. In this way, an immediate link between place and time adverbials could be established, as indicated in Quirk et al. (1985, pp. 479-482): "temporal relations are especially dependent for their expression upon figurative extension of locative items such as *in* and *at.*" Semantically, time is simple, and can be thought of as a sequence of points which are located on a time line. Instead of being multi-dimensional, "time is one-dimensional and has nothing analogous to the vertical axis (up-down) or the lateral axis (left-right). In addition, time is unidirectional in that for two points on the time line that do not coincide (i.e. are not simultaneous), one is unambiguously earlier and the other is later. Finally, time is not bounded on either side" (Haspelmath, 1997, p. 23).

What follows are the semantic roles of time adverbials distinguished in Quirk et al. (1985, pp. 479-482), some of which are similar to those of space: 1) 'position'. Time is seen as a fixed position on a scale; 2) 'duration'. Linking duration to specific positions on the linear time scale, we have the concept of 'span', which can be divided to two categories: the 'forward span' and the 'backward span'; 3) the expressions of 'frequency' can be elicited by 'How often . . .' ; 4) there is the expression of 'relationship' between one time and another. For example: *He had visited his mother already when I saw him yesterday*

3.2.3 Pronouns and pro-verbs (6-11)

In this section, six linguistic features from the category of pronouns and pro-verbs are reviewed: 6) First-person pronouns, 7) Second-person pronouns, 8) Third-person personal pronouns (excluding *it*), 9) Pronoun *it*, 10) Demonstrative pronouns (that, this, these, those as pronouns), 11) Indefinite pronouns.

In this category, several different types of pronouns are included, the grammatical functions of which are varied or somewhat overlapping. The use of both first and second pronouns indicates the involvement of the addresser and addressee respectively (Chafe, 1985). The first person pronoun is also a prominent feature in interactive communicative situations, or in some comparative works of written and spoken registers. Besides, third person pronouns are used with reference to people, and in the plural also to “inanimate” (e.g., Biber, 1986, pp. 384-414; Chafe & Danielewicz, 1987, pp. 106-107). Biber (1986, pp. 384-414) suggests that their co-occurrence with past tense verbs and perfect aspect verbs indicates a “narrative” function. In the current research, all three types of personal pronouns are included: first-person pronouns (e.g., *I, me, my, mine, myself; we, us, our, ours, ourselves*); second-person pronouns (e.g., *you, your, yours, yourself, yourselves*); third-person pronouns (e.g., *he, him, his, himself; she, her, hers, herself; its, itself; they, them*)

The pronoun *it* has a wide range of uses (e.g., Quirk et al., 1985, pp. 348-349; Quirk & Greenbaum, 1973). Apart from referring to inanimate objects as in the sample sentence 3.15, it also refers to non-count substances as in 3.16, or singular abstractions as in 3.17, and even to singular collections of people, as in 3.18. Besides, *it* could also be used as an “empty” or “prop” subject, especially in expressions denoting time, distance, or atmospheric conditions, as in 3.19; or an anticipatory subject in cleft sentences/in clauses with extra-position, as in the sample sentence 3.20.

3.15) He was Cardiff's founding store manager, opening the store in May 2006 and overseeing *it* until late 2011. [it=the store] (file patagonia20120426)

3.16) The first block gives you the feeling that you're about to read a very time consuming amount of information. You probably felt compelled to skip over *it* and read the second block instead, because it appeared easier to read.[it=information] (file INFSOFT20120619)

3.17) The most important thing to know about digital content is that people read *it* differently than print content. [it=digital content] (file corporateeye57)

3.18) Parliament's answer to all awkward problems is to establish a Royal Commission whose findings it can then ignore. (Quirk et al., 1985, p. 348)

3.19) What time is it? It's half past five. (Quirk et al., 1985, p. 349)

3.20) The bottom line with West Nile is that it's more of an issue for you and your human family that it is for the furry ones in your home. (file pu20120921)

Like the pronoun *it,* the demonstrative pronouns (*that, this, these, those as pronouns*) are often

used in speech and in more informal pieces of writing, without explicit reference (Chafe & Danielewicz, 1987, pp. 90-91). Besides, their use may also be considered under the headings of situational reference referring to the extra-linguistic situation as in the sample sentences 3.21 and 3.22, anaphoric reference referring to an earlier part of the discourse as in 3.23, and cataphoric reference referring to a later part of the discourse as in 3.24 (Quirk et al., 1985, p.372). And functionally, demonstrative pronouns are among the typical markers of informational spoken and online discourse (Biber, 1988, 1995; Collot & Belmore, 1996).

3.21) Earlier *this* month, nearly 7,500 people descended on Chicago, all driven by a common insatiable need to better understand the ever-evolving landscape of e-commerce. [more recently] (file Baderrutter37)

3.22) Later *that* month, the United Nations military operations started in Libya, prompting a surge in status updates mentioning 'Libya' and 'Gaddafi'.[some time before] (file Facebook20111206)

3.23) I hear you disliked his latest novel. 1 read his first novel, and *that* was boring, too. (Quirk et al., 1985, p.375)

3.24) He told the story like *this*: 'Once upon a time . . .' (Quirk et al., 1985, p.375)

In terms of indefinite pronouns (e.g., *anybody, nothing, someone*), they have been identified as being lacking in definiteness. This type of pronoun is dominated by the compound pronouns that are composed of two morphemes, a determiner morpheme, such as every-, some-, any-, or no-, and a nominal morpheme, such as -one, -body, or -thing. Taken together, this type of pronoun can be divided into four categories: 1) universal: everybody, everyone, everything; 2) assertive: somebody, someone, something; 3) nonassertive: anybody, anyone, anything; 4) negative: nobody, no one, nothing. This type of pronoun also includes another category called of-pronouns because they can be followed by the of-phrase: many (of), some (of), all (of), both (of), a few (of) etc. (Greenbaum & Quirk, 1990; Quirk et al., 1985). Functionally, they were included in Biber's (1988, p.226) study as markers of "generalized pronominal reference", which is the same case to pronoun *it* and the demonstrative pronouns. Kroch and Hindle (1982) have associated greater generalized pronoun use with the limited amounts of information that can be produced in a typical spoken situation.

3.2.4 Questions (12)

The language feature 'questions' has been either regarded as grammatical and semantic

categories (Crystal, 2011; Quirk et al., 1985) or as a discourse category (e.g., Burton, 1980; Tsui, 1991), thus being defined from different perspectives. According to Quirk et al. (1985, p. 804), questions could be defined as “a semantic class which is primarily to seek on a specific point”. There are three major classes of question according to the answer they expect: Yes/No questions are those “that expect affirmation or negation”, Wh-question are those “that typically expect a reply from an open range of replies”, and alternative questions are those “that expect as the reply one of two or more options presented in the question” (Quirk et al., 1985, p. 806).

In the current research, wh-questions has been selected as a linguistic feature for a multi-dimensional analysis. As its name show, wh-questions consist of the wh-words/ simple interrogative words, including *who/whom/whose, what, which, when, where, how* and *why*, as shown in examples from 3.25 to 3.32.

3.25) What are your earliest memories of surfing? (file patagonia20120502)
3.26) Which ice cream record should we attempt? (file 2turkeyhill)
3.27) When will you be promoted? (Quirk et al., 1985, p. 818)
3.28) Where shall I put the glasses? (Quirk et al., 1985, p. 818)
3.29) Who are you helping? (file FedEx20120504)
3.30) What are you trying to accomplish? (file FedEx20120504)
3.31) Why does it matter to you and your stakeholders? (file FedEx20120504)
3.32) How did you make conditions better for the recipient, whether a community, an ecosystem, or wildlife directly? (file FedEx20120504)

As a discourse category, questions have been characterized as a kind of request or directive (Burton, 1980), an elicitation to prospect a linguistic response (Tsui, 1994) and as the speech function realized by interrogatives (Halliday, 1994). Thus, semantically, for all different types of questions including wh-questions, they could be considered as a device to involve the reader in an interactive communication.

3.2.5 Nominal forms (13)

This section reviews a linguistic feature from the category of nominal forms, i.e., 13) total nouns. Semantically, nouns refer to concrete entities, such as people and things in the external world, as well as denoting qualities and states, e.g., freedom, friendship. They mainly play the role of “referential specification” in the discourse (Biber et al., 1999, p. 232). With a referential

use of nouns, we are able to know the people who are responsible for the actions in a series of events, descriptions, or debates/arguments. When each participant role is constructed in a nominal group, it can not only contribute different kinds of meanings, but also refer to something in the situational context. In some previous studies (Biber, 1988; Rowley-Jolivet & Carter-Thomas, 2005; Westin, 2002a), it has been agreed that nouns are used to compact information.

3.2.6 Passives (14-15)

This section includes the review of two linguistic features: 14) agentless passives and 15) by-passives. Passives of some verbs have been identified as a typical, or the most common characteristic of scientific discourse (e.g., Banks, 2008; Biber et al., 1999). In previous studies, the preference for passives in academic/ scientific discourse has been motivated by discourse needs (Hyland, 1996) in general, and associated with the need to obscure the author's identity (Biber, 2012; Dorgeloh & Wanner, 2009), or the need to mention the agent (Nesi, 2011) in particular.

Passives could be classified into two categories: passive with agent (by-passives), which is also called the long passive, and passive without agent (agentless passives), which is called the short passive (Quirk et al., 1972; Quirk et al., 1985, p. 167). In terms of agent roles, agentless passivesare used mostly when the role of the agent is not that important or unknown, as in example 3.33. In terms of the choice of by-passives, the reasons could be an intention to put the emphasis on agent rather than the action, as in example 3.34, or to avoid long subjects (Greenbaum and Quirk, 1993), as in 3.35. Both of by-passives and agentless passives have the same function: to switch the focus from subject to object in the active clause or action itself.

3.33) Frankly, most organizations can benefit from learning, given that real needs of a specific organization are identified so that the most appropriate solutions *can be implemented*. (file UpsideLearning20120816)

3.34) Her first book- The Story of Stuff- *was published by* Free Press in March 2010. (file patagonia20120412)

3.35) From the other side of the Israeli political spectrum, Arutz Sheva, a news site *run by Israeli settlers who live in the disputed settlements built in the West Bank*, reported on Mr. Adnan's appeal to Israel's High Court to intervene under the headline, "Hunger Striking

Terrorist Headed to Supreme Court." (file nytimes2-20120216)

3.2.7 Stative forms (16)

This section includes a review and discussion on the linguistic features 'existential *there*' from the category of stative forms. Early linguists, such as Fillmore (1968), Anderson (1971), etc., have commonly agreed that *there* is a type of locative adverbial referring to location. More recently, Biber et al. (1999, p. 944) have further explicated the grammatical status of existential *there* in detail: "the original locative meaning in it has lost, and syntactically, it functions as a grammatical subject rather than as an adverbial".

Typically, existential *there* is used together with an intransitive verb *be,* as shown in the following structure: there + be + indefinite NP (+ place or time position adverbial). The form of the verb *be* could be varied, apart from simple be (e.g. is, are, was, were): it could be preceded by auxiliaries or semi-modals, such as *will be*, *is to be*, *is supposed to be*, *used to be,* as shown in 3.36; or appears in a to-infinitive complement of a lexical verb with the force of a hedge, e.g. *happen to be*, *tend to be*, *appear to be*, as in 3.37. Besides, existential clauses may also contain verbs other than *be*, chiefly intransitive verbs denoting existence or occurrence, such as *exist*, *come*, *seem*, etc., as in 3.38. Thus, it is shown that existential *there* by no means focuses on the existence or occurrence of something, or in other words, "introduces new elements into the discourse" (Biber, 1999, p. 951). However, they are slightly used in different contexts. One context is focusing on the existence of something that is put at the beginning of a spoken/written narrative, which is then picked up by later references in the text, as in 3.39. Another way of using existential *there* is to use it to introduce a series or even a sequence of elements, as in 3.40.

3.36) *There will be* extra costs for the replication function in the storage and FC connection, but this may be less important than the recovery time for the application and in the end be more efficient. (file HITACHI20120410)

3.37) *There appears to be* agreement that acceptable lead-free reliability has been established for consumer products with lifetimes of 5 years or so, but not for military/aerospace electronics where lifetimes can be up to 40 years in harsh service conditions. (file Indium20111212)

3.38) *There seemed to be* uniform agreement that solder paste users should be able to demand that their lead-free solder paste perform well with any PWB pad finish (e.g. OSP

Immersion silver, electroless nickel gold, etc.) without the use of nitrogen. (file Indium20111212)

3.39) *There are* many examples of how big data brings value to companies large and small. Collecting data from new sources and correlating the data to generate information that helps grow the business and provide a competitive edge. I will be providing several examples of this during my series on big data. However, the value of big data extends beyond commerce. (file HITACHI20120529)

3.40) Instead of stencil printers and component placement machines, *there was* an MRI machine. *There were* techs that ran the MRI machines just like there were operators. The nurses were like the process engineers, and *there were* some medical doctors that were like the mangers and execs at her company. (file Indium20120120)

3.2.8 Subordination features (17-24)

In this section, eight linguistic features from the category of subordination features are reviewed: 17) infinitives, 18) THAT relative clauses controlled by a verb, 19) THAT relative clauses controlled by an adjective, 20) THAT relative clauses controlled by a verb, 21) WH relative clauses, 22) concessive adverbial subordinators (although, though), 23) causative adverbial subordinators (because), 24) conditional adverbial subordinators (if, unless). All of these linguistic features have a wide range of different discourse functions, as discussed in the following paragraphs.

According to Biber et al. (1999, p. 693), infinitives are used to "report intentions, desires, efforts, perceptual states, and various other general actions". All these different functions are realized through different grammatical structures occurring in either post-predicate or subject position (Biber et al., 1999, pp. 693-740), which, in specific, include "post-predicate infinitive clauses controlled by verbs", as in 3.41, "verbs taking extra posed to-clauses", as in 3.42, "subject noun phrases and subject predicative", as in 3.43, and "to-clauses controlled by adjectives", as in 3.44.

3.41) We'll always *try to* add a heaping spoonful of reality, and a dash of zest. (file AVAYA20120207)

3.42) I love going to trade shows, it's a great way *to interact* with those who use the products and hear their reactions. (file Fiskateers114)

3.43) The aim of EarthSmart is *to improve* the company's environmental performance and find ways for FedEx to contribute to sustainability on a broader scale by working with others who share that vision, both within and out of the Company. (file FedEx20120417)

3.44) Rose has the Rubbermaid Configurations line of closet in her daughter's room so it

was easy *to install* extra shelves and will be easy *to rearrange* or take them out as her daughter grows. (file rub2)

Among the varied structures of infinitives, the most commonly used are those in verb and adjective complementation, as in 3.45 and 3.46, encoding the speaker's attitude. Thus, it is considered to be reasonable to include infinitives among the markers of argumentative discourse (Biber, 1988). Besides, infinitive clauses are also considerably more common in the written registers than conversation. This is in marked contrast to the register distribution of that-clauses and wh-clauses, which are both most common in conversation (Biber, 1999, p.699).

3.45) The good news is that professionals *want to* join these boards. According to the Taproot Foundation, an estimated 87% of human resources professionals and 92% of marketing professionals are interested in board service but far fewer actually serve on boards. (file LinkedIn20120917)

3.46) Since the G3X and area 79X each have two serial ports, it is rather simple *to interface* these units to both a GDL 39 and external serial device. (file GARMIN20120808)

Both of that-clauses and wh-clauses, particularly those included in the current study, i.e. 18) THAT relative clauses controlled by a verb, 19) THAT relative clauses controlled by an adjective, 20) THAT relative clauses controlled by a verb, and 21) WH relative clauses, can be categorized into 'relative clauses'. However, they have different discourse functions depending on their positions in the sentence: If the subordinate clause is placed at the beginning of the sentence, it can play an important role in establishing a cohesive and discourse-organizing link between the text preceding and after the clause. On the other hand, when the adverbial clause is placed at the end of the sentence, it provides expansion of the information in the main clause (Celce-Murcia, 1998; DeCarrico, 2000). Depending on how a subordinate clause is linked to the main clause, several major categories of relative clauses could be distinguished, as listed below:

1) The adnominal relative clause

The adnominal relative clause is the central type of relative clause. This type of relative clauses forms part of a constituent of the matrix clause, in particular, "post-modifying noun phrases with any kind of lexical meaning whose relative markers can serve different sorts of grammatical functions in the relative clauses (subject, object, complement, or adverbial)"

(Quirk et al. 1985, p.1257-59). Within adnominal relative clauses, two subtypes can be distinguished at semantic or pragmatic levels (Biber et al. 1999, p. 195): “restrictive relative clauses that are used to establish the reference of the antecedent”, as in 3.47, and “non-restrictive relative clauses used to present additional information which is not required for identification”, as in 3.48.

3.47) We have 30 men who are working from 6am to 11 p m and most of the extra payments we would expect to receive may go on overtime. (Biber et al. 1999, p. 195)

3.48) He warned the public not to approach the men, who are armed and dangerous.(Biber et al. 1999, p.195)

In terms of the distribution of the two sub-types of adnominal relative clauses, “restrictive relative clauses are much more common than non-restrictive clauses (marked by a comma) in all written registers” (Biber et al., 1999, p. 603). The distribution of the two sub-types varies somewhat by register: non-restrictive relative clauses take up a smaller percentage of all relative clauses in fiction, academic prose, and news.

2) The nominal relative clauses

This type of clause is unique among relative clauses in that they form a constituent of the matrix clause, by means of containing their antecedents. As defined by Quirk et al. (1985, p. 1056-61), the nominal relative clause is “basically a noun phrase modified by an adnominal relative clause, except that its wh-element is merged with its antecedent (i.e., the phrase to which the wh-element refers)”. That is, the antecedent is incorporated in the relative marker. Such a relationship can be demonstrated in example 3.49 and 3.40 (Quirk et al., 1985, p. 1056).

3.49) I eat what I like.

3.50) I eat that which I like.

3) The sentential relative clauses

Unlike nominal and adnominal relative clauses, sentential relative clauses do not constitute an element of the matrix clause. That means, rather than being nominalized, they are sententialized, or clausalized. They post modify other phrases or larger chunks than mere noun phrases, referring back to the predicate or predication of a clause, as in 3.51 and 3.52, or to a whole clause or sentence, as in 3.53 and 3.54, or even to a series of sentences, as in 3.55 (Quirk et al., 1985, pp. 1118-1120).

3.51) They say he plays truant, which he doesn’t.

3.52) He walks for an hour each morning, which would bore me.
3.53) Things then improved, which surprised me.
3.54) Colin married my sister and I married his brother, which makes Colin and me double in-laws.
3.55) - which is how the kangaroo came to have a pouch.

Another set of linguistic features in this category includes different types of subordinators: 22) concessive adverbial subordinators (although, though), 23) causative adverbial subordinators (because), 24) conditional adverbial subordinators (if, unless), which have been selected for the MD study in Biber's 1988 study. The term "subordinator" has been defined as "subordinating conjunctions, words which introduce (mainly finite) dependent clauses" (Biber et al., 1999, p. 85). Functionally, the same subordinators may introduce various types of adverbial clauses, and present clauses with various meaning (Quirk et al., 1985, p. 1077). In general, the differences among various approaches and views on the semantic classification of adverbial clauses are not so essential. An important example is the clauses of time, place, reason and condition are commonly used by many researchers, such as Biber et al. (1999, pp. 818-819), Eastwood (1994, p. 327) and Close (1978, pp. 56-63). In spite of this, different semantic roles as well as syntactic functions of adverbial clauses have been identified and presented by Biber (1988). The syntactic functions of adverbial clauses which include "circumstance adverbials (adjuncts), stance adverbials (disjuncts) and linking adverbials (conjuncts)" in Biber et al. (1999, p. 763) are considered to be the same as those in Quirk (1985). Presented below are 3 semantic roles of adverbial clauses.

1) concessive adverbial subordinators (although, though)

A concessive clause implies the contrast between the main and subordinate clause, and the situation that the main clause is contrary to the expectation of what is said in a concessive clause (Quirk et al., 1985, pp. 1097-1098). Concessive clauses are mainly introduced by *although* or more informal variant *though*. Other subordinators could be *even though*, *whereas*, *while* etc. The concessive subordinators may introduce –ing, -ed and verbless clauses, as in 3.56 and 3.57 (Quirk et al, 1985, p. 1097):

3.56) Though well over eighty, he can walk faster than I can.
3.57) While not wanting to seem obstinate, I insisted on a definite reply.

2) causative adverbial subordinators (because)

Alexander (1988, p. 26) claims that causative adverbial subordinators answer the question "why?" and could be introduced by *because*, *as* or *since*. Quirk et al. (1985, pp. 1105-1106) add that they could also be introduced by prepositional phrases, such as "because of, due to, owing to, on account of", etc.; however, these prepositional phrases are considered to be clumsy due to their length. These subordinators are often named "adjuncts"; yet, they are "disjuncts" when they are positioned finally (Quirk & Greenbaum, 1973, p. 327). Greenbaum and Quirk (1990) hold the viewpoint that a reason clause answers the question 'why' and may express either a direct or an indirect reason. A direct reason conveys a cause and effect, reason and consequence, motivation and result, or circumstances and consequence relationship (Quirk et al., 1985, pp. 1103-1104) . Most subordinators in direct clause introduce "content disjuncts" (Quirk et al., 1985, p. 1106), as can be seen in the following sample sentences (Quirk et al., 1985, pp. 1103-1104):

3.58) The flowers are growing so well because I sprayed them.
3.59) She watered the flowers because they were dry.
3.60) I watered the flowers because my parents told me to do so.
3.61) Since the weather has improved, the game will be held as planned.

Besides, in comparison with a direct reason, the use of an indirect reason is more peripheral and less related to the situation in the main clause. 'Since' and 'because' are the most common subordinators for an indirect clause and their function is usually a "style disjunct" (Quirk et al., 1985, p. 1104). A sample sentence of an indirect reason is presented below:

3.62) Vanessa is your favourite aunt, because your parents told me so.(Quirk et al, 1985, p. 1104)

3) conditional adverbial subordinators (if, unless)

Conditional clause expresses a cause-effect relationship and it is about the possibility of the happening of an event that can or might occur or might have occurred. Eastwood (1994) adds that conditional clauses are used when people want to advice, criticize, offer, request etc. Quirk et al. (1985, pp. 1088-1091) state that conditional clauses may express either direct or indirect condition. A direct condition conveys a direct situation, which may be either open or hypothetical: open conditions are neutral and do not answer the question of fulfillment of the condition, while a hypothetical condition conveys the speaker's belief of the unfulfillment of the

condition. In the following sample sentences, 3.63 presents an open condition, and the conditional clauses in 3.64 and 3.65 present a hypothetical condition.

3.63) Read the paper if you don't believe me! (Biber et al, 1999, p. 819)

3.64) They would be here with us if they had the time. (Quirk et al, 1985, p. 1091)

3.65) I would if I was you. (Biber et al, 1999, p. 819)

The distinction between an open and a hypothetical condition is important because “verbs in hypothetical conditions are back-shifted” (Quirk et al, 1985, p. 1091).

In comparison with the direct condition, an indirect condition is less frequently used and is not related to the situation in the main clause. Direct if-clauses function as adjuncts and indirect conditions function as style disjuncts (Quirk et al, 1985, p.1089). Eastwood (1994) on the other hand divides conditional clauses into open and unreal conditions according to the degree of reality expressed. The difference between a direct and an indirect condition together with the syntactic function and form is shown on the following sample sentences (Quirk et al, 1985, p. 1089):

3.66) If you put the baby down, she’ll scream. (a direct condition that functions as an adjunct)

3.67) She’s far too considerate, if I may say so. (indirect condition dependent on the utterance)

To sum up, a conditional clause expresses cause and a consequence relationship and may be either direct (that may be open or hypothetical) or indirect. Direct conditional clauses function as adjuncts whereas indirect conditional clauses as style disjuncts.

3.2.9 Prepositional phrases, adjectives, and adverbs (25-27)

In this section, three linguistic features from the category of are reviewed briefly: (25) total prepositional phrases, (26) total adjectives, (27) total adverbs. Among the three linguistic features, prepositional phrases and adverbs are the syntactic structures that realize the linking adverbials. According to Biber (1999, p. 862-863), “single adverbs can be used to convey virtually every kind of stance meaning with respect to a proposition”, including certain vs. doubt, actuality vs. reality, evidence, particularly with apparently and evidently, limitation, etc. Single adverbs are also used as attitude adverbials, conveying an attitude towards a proposition or a judgement. Besides, single adverbs can also be used to communicate in a casual style. In his

study, it has been found that almost all linking adverbials are single adverbs in conversation, while prepositional phrases are also relatively common as linking adverbials in academic prose.

The prepositional phrase, as a language feature, is closely related to informational discourse (Chafe, 1985). It has been found by Chafe (1982) and Chafe and Danielewicz (1987) that prepositional phrases are frequently used in all the genres they studied, such as conversations, lectures, personal letters, and particularly frequent in academic writing. Biber's 1988 (236-237) study also proved this and showed that the prepositional phrase indicates information density. Similarly, adjectives also contribute to the information density of a text (Chafe, 1982; Chafe and Danielewicz, 1987).

3.2.10 Lexical specificity (28-29)

In this section, two linguistic features from the category of lexical specificity are reviewed: (28) Type-token ratio, (29) Mean word length. Both of the two features are among the most commonly used measures of lexical specificity and diversity, which are also considered to be associated with information density, and production circumstances (Biber, 1988; Chafe & Danielewicz, 1987; Zipf, 1949).

According to Biber (1988, p.238), type-token ratio is "the number of different lexical items in a text, as percentage". It is one way to measure lexical specificity, which reflects vocabulary range, the richness of the vocabulary of a text, and also the density of information. It is calculated by dividing the number of different words appearing in the text (types) by the total number of words in the text (tokens). A high type/token ratio in a text indicates a more varied and extensive use of vocabulary. As a measurement of vocabulary range, and information density, it has been revealed by many researchers (Biber, 1988; Chafe & Danielewicz, 1987; Zipf, 1949) that high frequencies of lengthy words, and type/token ratio, as well as prepositional phrases, attributive adjectives, and nouns, indicate high density of information integrated into a text. Besides, longer words also convey more specific, specialized meanings than shorter words (Zipf, 1949).

In previous studies, the use of TTR has been explored in different registers by both linguistic and non-linguistic researchers. In the 1960s and 1970s, particularly, TTR were frequently used by psychologists and researchers in their study of differences between spoken and written language. For example, Chafe and Danielewicz (1987, p.88) have found that the TTR was higher in the written genres than in the spoken ones due to their differences in production circumstances. It is a very difficult task to achieve a precise lexical choice and production in spoken genres. Besides, the use of TTR has also been studied in registers other than spoken and written languages. For example, Tagliacozzo (1976) has proved the hypothesis that the type-token ratio (TTR) of scientific and technical writings varies in relation to the degree of technicality of the text. More specifically, highly technical writings have lower TTRs than writings of lower technicality. In a very recent study, Johansson (2012) has found that TTR is used differently in the four main registers. It is consistently lower in conversation than in the written registers, and somewhat lower in academic prose than in fiction and news reportage. The suggested explanation is that TTR is low in conversation because it is less concerned with the transmission of information than writing, and its production reflects a more spontaneous process which requires a little time for planning and varying the choice of words. Besides, conversation involves more repetition, which may be used for emphasis, to help the planning of the speaker, or to make sure that the message gets across to the hearer. On the other hand, the TTR differences among the written registers are even more surprising. Instead of having a higher TTR in fiction, the TTR in news is very high which reflects the extremely high density of nominal elements in that register, referring to a diverse range of people, places, objects, events, etc. At the other extreme, academic prose has the second lowest TTR, reflecting the fact that a great deal of academic writing has a restricted technical vocabulary and is therefore less variable than fiction and news report.

Another linguistic feature in this category is mean word length, which is also a way of measuring lexical specificity and diversity. Klare (1963) has traced the modern history of readability studies to the 1920s when many readability formulas were introduced by educational researchers, and provided a possible interpretation for how sentence length works

in predicting passage difficulty. In terms of the mean word length and spoken/written language, Chafe and Danielewicz (1987, p. 88) suggest that spoken language has more general, shorter and fragmented words, while the longer and more specific words are normally found in written texts. Some CMC researchers (e.g., Murray, 1990; Werry, 1996) attribute the use of spoken language in CMC to the restrictions of on-line production, which does not often provide the author/speaker enough time to choose their words. Besides, in Biber's (1986) study of co-occurrence of certain language features in the oral and written discourse of native speakers of English, he found that longer words convey more specialized meanings than shorter ones. He also claimed that a high average word length indicates a "highly exact presentation of informational content in a text, conveying a maximum content in the fewest words" (Biber, 1986, p. 394).

3.2.11 Modals (20-32)

In this section, three categories of modal verbs are discussed: (30) Possibility modals (can, may, might, could), (31) Necessity modals (ought, should, must), (32) Predictive modals (will, would, shall) (Biber, 1988, p.241). This classification is on the basis of their function in language: predictive modals for referring to the future; possibility modals for linguistically realizing different perspectives on a topic; and necessity modals, which directly express persuasiveness.

Each modal can have two different types of meanings, which can be labelled intrinsic and extrinsic: intrinsic modality refers to "actions and events that humans directly control, the meanings of which relate to permission, obligation, or intention", and extrinsic modality refers to "the logical status of events or states, usually relating to assessments of likelihood: possibility, necessity, or prediction" (Biber et al., 1999, pp. 485-486; Quirk et al., 1985, pp. 219-221).

As far as possibility modals (can, may, might, could) are concerned, their meanings and uses are varied to differing extents, but follow some principles (Biber et al., 1999, pp. 485-486; Quirk et al., 1985, pp. 219-224). Firstly, some modals are used only to mark one meaning; for

example, *might* is used only to mark logical possibility, as in 3.68. Secondly, some of them have both intrinsic and extrinsic uses: for example, *may* has the meaning of permission (intrinsic), as in 3.69, and the meaning of possibility (extrinsic), as in 3.70; *can* commonly marks permission, ability, and logical possibility, as in 3.71, 3.72, and 3.73; *could* in conversation is also sometimes used to mark ability or permission, as in 3.74 and 3.75; Thirdly, some modals tend to have overlapping/ambiguous meanings, for example, both *can* and *could* ambiguously mark logical possibility or ability (or permission), as in 3.76 and 3.77.

3.68) It *might* be better for airlines to be more realistic and expect people to bring their own meals. (file IAG20120803)

3.69) You *may* borrow my bicycle if you wish. (Quirk et al., 1985, p. 223)

3.70) It's important that your social media presence is a good one because customers *may* find that before they find your actual website. For example, a Tweet that contains your company name *may* rank higher than your company blog. (file chr20120712)

3.71) *Can* we borrow these books from the library? (Quirk et al., 1985, p. 222)

3.72) As Jakob Neilsen points out, currently native apps are the best in delivering top notch experience, if you *can* afford it. (fileUpsideLearning20120816)

3.73) When you are talking about the success or failure of your marketing program, the proper collaborations *can* mean everything. (file chr20120710)

3.74) When I was a kid, I read every book I *could* find on the South Pacific, and this is what I wanted to do. (file Patagonia20120514)

3.75) Walk around your vehicle before you leave to ensure there are no children or objects that you *could* hit. (file FedEx20120507)

3.76) You can read more about the San Francisco 49 success story here. (file HITACHI20120618)

3.77) I wish I *could* single out everyone who deserves it. But I'd love if you *could* share any stories about exceptional associates you've met in our hotels. (file Marriott20120509)

In terms of necessity modals (ought, should, must), they also have both intrinsic and extrinsic uses. In particular, *must* is used commonly for both logical necessity (extrinsic) and personal obligation (intrinsic), as in 3.78 and 3.79. The other two modals *should* and *ought to* are also used to mark personal obligation (rather than logical necessity), as in 3.80 and 3.81. Besides, two additional meanings of *should* can be distinguished: a first person alternative of hypothetical *would*, as in 3.82; and a marker of the putative meaning, as in 3.83 and 3.84 (Quirk et al., 1985, p. 234).

3.78) The face *must* be bigger than we thought, since we'd go down at least a thousand feet from the summit of the ca 6,200' peak before starting up San Valentin. (file patagonia20120413)

3.79) Throughout my adult years of taking vacations I have realized that there are some very important things I *must* do to keep myself sane both before I leave and after I return. (file rub1)

3.80) I am finally doing what I said that you *should* do: listen 80% and talk 20% (who has the time, right?). (file chr20120810)

3.81) The same year they launched a voluntary program called Project Gutenberg to make electronic copies of cultural texts that *ought to* be placed in the archives. (file Business Property20120214)

3.82) If there were an accident, we would/should have to report it.(Quirk et al., 1985, p. 234)

3.83) It's unfair that so many people should lose their jobs.(Quirk et al., 1985, p. 234)

3.84) Why should anyone object to her enjoying herself?(Quirk et al., 1985, p. 234)

As far as predictive modals (*will, would, shall*) are concerned, all of them can be used to mark both volition (intrinsic) and prediction (extrinsic) (Quirk et al., 1985, pp.228-231; Biber et al., 1999, pp. 495-497). Under the heading of 'prediction', three related uses of *will/would* can be distinguished.

i. The common future predictive sense of *will* corresponds to prediction in the past sense of *would*, as in 3.85 and 3.86.

3.85) This expansion *will* create new jobs and put more Kentuckians back to work. (file toyota20120508)

3.86) I was told I *would* feel better after this medicine. (Quirk et al., 1985, p. 228)

ii. The present predictive sense of *will* is comparatively rare and similar in meaning to *must* in the 'logical necessity' sense, as in 3.87.

3.87) She will have had her dinner by now. (Quirk et al., 1985, p. 228)

iii. The habitual predictive meaning which often occurs in conditional sentences, as in 3.88, or timeless statements of predictability, as in 3.89.

3.88) I have included a video that *will* show you the machine and explain how it works. (file Xerox20120911-2)

3.89) Oil dill float on water.(Quirk et al., 1985, p. 228)

Besides, three related uses of will/would can also be distinguished under the heading of willingness, which ranges from the weak to the strong willing, as in 3.90 and 3.91. Between these two, there is the more usual sense of intention, which often combines with a sense of prediction, as in 3.92. In comparison with *will/would, shall* is generally used to mark prediction, as in 3.93; or personal intention, as in 3.94 and 3.95. In 3.94, *shall* is again a formal (and

traditionally prescribed) alternative to *will* in an intentional sense. In 3.95, a question is contained which consults the wishes of the addressee, and thus moves from an intentional towards an obligational meaning.

3.90) I have many ideas that might help; let me know if you are interested and I *will* send you some. (file Xerox20120917)

3.91) As long as media keep offering PR pros the chance to advance important issues, I *will* stay determined to chase after the stories and campaigns that deliver these life-changing headlines.(file Brand Channeler20120309)

3.92) *Will* you be checking out this Jackson documentary during the Thanksgiving holiday? Meanwhile, check out some promo photos below. (Fandango20120829-12)

3.93) According to the opinion polls, 1 shall win quite easily. (Quirk et al., 1985, p. 230)

3.94) We shall uphold the wishes of the people. (Quirk et al., 1985, p. 230)

3.95) By the way, construction workers and farmers are also living a harsh life in China, *shall* we also boycott housing and grains? (file nytimes20-20120125)

3.3 Summary

This chapter provides a detailed description of the linguistic features selected for the MD analysis. In the description of each of them, it is not only the grammatical and semantic roles, but also the discourse functions that have been introduced. According to their shared linguistic functions, the linguistic features could be classified into several categories, as presented below:

1. Indicating personal involvement - 3. Present tense; 6. First-person pronouns; 7. Second-person pronouns; 9. Pronoun it; 10. Demonstrative pronouns (that, this, these, those as pronouns); 11. Indefinite pronouns (e.g., anybody, nothing, someone; 12. Direct WH questions; 30. Possibility modals (can, may, might, could);
2. Indicating information density – 4. Place adverbials (e.g., above, beside, outdoors); 5 Time adverbials (e.g., early, instantly, soon); 13. Total nouns; 25. Total prepositional phrases; 26. Total adjectives; 27. Total adverbs; 28. Type-token ratio; 29. Mean word length;
3. Indicating narrative discourse – 1. Past tense; 2. Perfect aspect; 8. Third-person personal pronouns (excluding it); 16. Existential there;
4. Indicating argumentative discourse - 22. Concessive adverbial subordinators (although, though); 23. Causative adverbial subordinators (because); 24. Conditional adverbial subordinators (if, unless); 31. Necessity modals (ought, should, must); 32. Predictive

modals (will, would, shall)

5. Indicating abstract discourse - 14. Agentless passives; 15. by-passives;
6. Indicating reference - 18. THAT relative clauses controlled by a verb; 19. THAT relative clauses controlled by an adjective; 20. THAT relative clauses controlled by a verb; 21. WH relative clauses.

This typology of linguistic functions encompassing various linguistic features shed light on the interpretation of factors and clusters identified in the factor and cluster analysis, which are the two component parts of the MD analysis. The results of MD analysis will be presented in detail in Chapter 5. Prior to this, the issue of corpus design and methodology is discussed in Chapter 4.

Chapter 4: Corpus design and methodology

4.1 Introduction

This chapter addresses some issues directly related to the research design that I adopt in this research project. In order to be well suited towards my research goals and the conceptual framework, the issues concerning corpus design, methodology, and data processing are to be addressed. This chapter begins with a brief description of the corporate blog corpus (CBC) used for the current research (Section 4.2). This is followed by a discussion of some basic corpus design issues underlying the creation of CBC, which include the aim and the type of the corpus, representativeness, corpus size and sample size, data collection and ethical issues (Section 4.3). Subsequently, the methodological approaches employed in the current research are introduced (Section 4.4), which in general include Biber's (e.g., 1988, 1995) MD analysis to register variation, and SFL-based approach (e.g., Halliday & Matthiessen, 2004) for the exploration of salient linguistic differences across different CB text types. In particular, several major issues in MD analysis are introduced and discussed, including selection of linguistic features (Section 4.4.1.1), text tagging and frequency counts of linguistic features (Section 4.4.1.2), factor analysis (Section 4.4.1.3), determination of the number of factors to retain (Section 4.4.1.4), cluster analysis (Section 4.4.1.5), and the exploration of CB text types across industries (Section 4.4.1.6). After this, how data is processed for the exploration of lexico-grammatical features in CB by using SFL-based approaches are introduced; the language features include the most common verbs and their process types in Section 4.4.2.1, personal pronouns in Section 4.4.2.2, modal verbs in Section 4.4.2.3, lexical density in Section 4.4.2.4, and grammatical complexity in Section 4.4.2.5.

4.2 Corpus description

CBC is a 590,520-word corpus consisting of 1,020 texts (named CB texts afterwards), which were retrieved from 41 top-ranked blogs (among top 50) listed on BlogRank (http://www.blogmetrics.org/Corporate_Blogs), a publicly available blog tracking and evaluating website. The selection of BlogRank is due to its availability as well as its authorities

in this area; it provides lists of “influential” and/or “top-ranked” corporate blogs or ratings of blogs (Torres-Salinas, Cabezas-Clavijo, Ruiz-Pérez, & López-Cózar, 2011). In BlogRank, over 20 factors, such as Yahoo incoming links, Google PageRank, unique monthly visitors, pages per visit, Technorati ranking, link to page ratio, etc., are used to rank the corporate blogs in different categories, each of which includes different numbers of corporate blog titles. All the corporate blog titles are linked to their webpages. Presented below is Figure 4.1, which is a webpage from BlogRank showing different factors to rank the blogs.

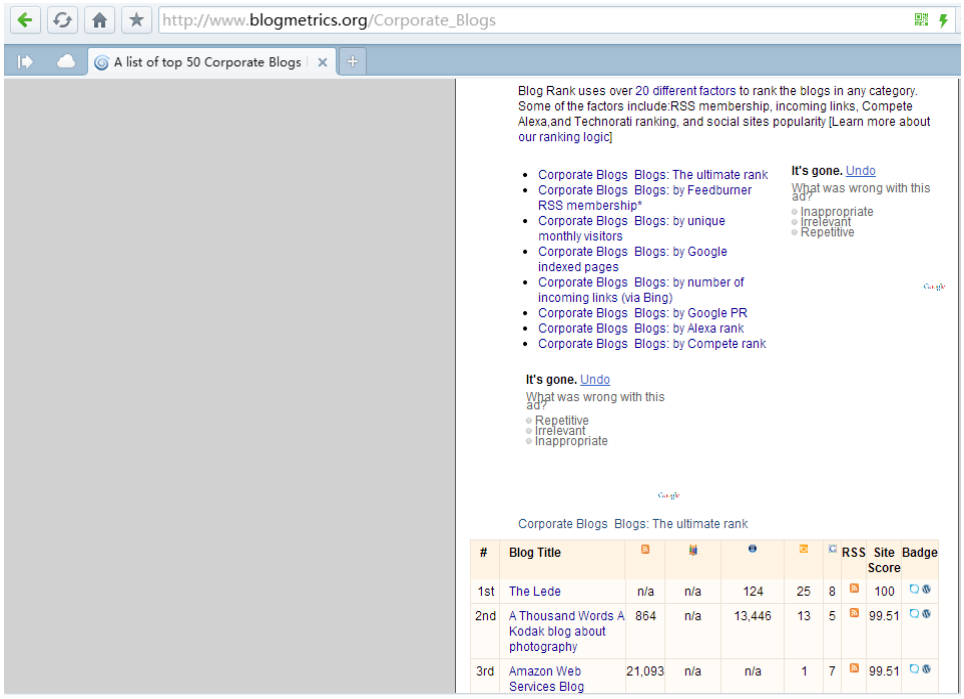


Figure 4.1: A webpage from BlogRank showing different factors to rank the blogs-1 (http://www.blogmetrics.org/Corporate_Blogs#ultimate)

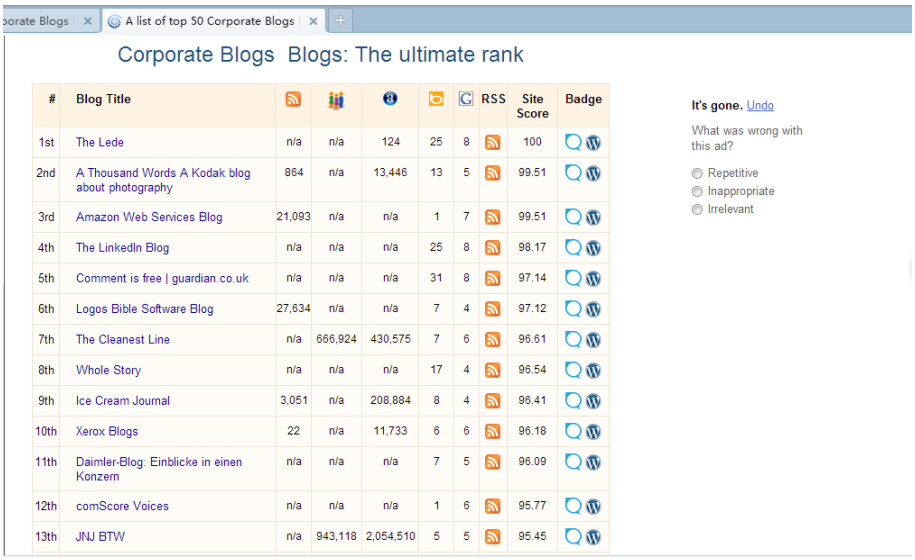
In the current research, it is decided that the corporate blogs by the ultimate rank could be used for the design of CBC, which is assumed to be a more comprehensive factor for blog ranking. A detailed description showing both the construction and the specification of the CBC is provided below. Besides, a complete list of the titles of originally retrieved 1058 CB texts, which includes 1020 CB texts in the CBC and other 38 texts excluded in the corpus, is provided in Appendix 1.

Table 4.1: The construction of CBC

No. of corporate blogs	41
No. of texts	1020
No. of words	590,520
Collection period	May 2012 ~ August 2012

Names of corporate blogs & Text count	1) Delta Air Lines Blog	25
	2) Whole Story	25
	3) Marriott on the move	30
	4) The Cleanest Line	22
	5) Toyota open road blog	27
	6) Red Cross Chat	28
	7) FedEx Citizenship Blog	30
	8) Facebook Blog	20
	9) Turkey Hill	29
	10) Lede	20
	11) Wallet	5
	12) Comment is free | guardian.co.uk	7
	13) Amazon Web Services Blog	25
	14) Avaya - The Blog	25
	15) Brand Channeler	25
	16) BreakingPoint Labs Blog	25
	17) Business Property	25
	18) Clorox - Dr. Laundry	25
	19) ComScore Voice	22
	20) Converge - A Bader Rutter Blog	23
	21) Corporate eye	22
	22) Dr. Laskys Blog	19
	23) Fiskars Ambassador	27
	24) Fandango Movie Blog	28
	25) Garmin Blog	38
	26) Hu Yoshida	54
	27) IAGblog	29
	28) Infusionsoft Blog	21
	29) JNJ BTW	20
	30) Logos Bible Software Blog	30
	31) Marketing Conversation	21
	32) Oracle Blogs	22
	33) Rubbermaid - Adventures in Organization	20
	34) RS Consulting blog	35
	35) PurinaCare Pet Insurance Blog	19
	36) The LinkedIn Blog	28
	37) Tyson Foods Hunger Relief	20
	38) The Unboxed Thoughts Blog	25
	39) Upside Learning Blog	30
	40) Verizon-PolicyBlog	25
	41) Xerox Blogs	23

As can be seen in Table 4.1, CBC consists of 41 corporate blogs, each of which includes different numbers of CB texts. There are totally 1020 CB texts with 590,520 words in the corpus, which was collected from May 2012 to August 2012. As mentioned earlier, the creation of this corpus relies on the internet resources, i.e., BlogRank (http://www.blogmetrics.org/Corporate_Blogs), a publicly available blog tracking and evaluating website where the blogs are listed in one place. And each of these corporate blogs is linked to their homepage (Figure 4.2).



#	Blog Title						RSS	Site Score	Badge
1st	The Lede	n/a	n/a	124	25	8		100	
2nd	A Thousand Words A Kodak blog about photography	864	n/a	13,446	13	5		99.51	
3rd	Amazon Web Services Blog	21,093	n/a	n/a	1	7		99.51	
4th	The LinkedIn Blog	n/a	n/a	n/a	25	8		98.17	
5th	Comment is free | guardian.co.uk	n/a	n/a	n/a	31	8		97.14	
6th	Logos Bible Software Blog	27,634	n/a	n/a	7	4		97.12	
7th	The Cleanest Line	n/a	666,924	430,575	7	6		96.61	
8th	Whole Story	n/a	n/a	n/a	17	4		96.54	
9th	Ice Cream Journal	3,051	n/a	208,884	8	4		96.41	
10th	Xerox Blogs	22	n/a	11,733	6	6		96.18	
11th	Daimler-Blog: Einblicke in einen Konzern	n/a	n/a	n/a	7	5		96.09	
12th	comScore Voices	n/a	n/a	n/a	1	6		95.77	
13th	JNJ BTW	n/a	943,118	2,054,510	5	5		95.45	

Figure 4.2: A webpage from BlogRank showing a list of ranked corporate blogs-2 (http://www.blogmetrics.org/Corporate_Blogs#ultimate)

Considering that a corporate blog title is too vague and does not provide enough background information of the blogs, a brief introduction to the 41 corporate blogs will be given below. This introduction plays an essential role in laying a foundation for organizing companies into industrial groupings in the later MD analysis.

1) **Delta Air Lines Blog**: This is a blog for Delta, one of the world's largest global airlines, helping more than 160 million travelers get to the places they want to go to each year. Delta's transportation strategies have been a key to their success, such as joint ventures and alliances with other airlines, airport expansions, increased flight offerings to new locations; more flights increasing seat capacity, etc. (http://www.blog. delta.com retrieved in 2012).

2) **Whole Story**: Whole Story is the official blog of Whole Foods Market, the world's leading

natural and organic grocer selling healthy food. Whole Foods Market's organizational structure supports business profitability through sales talk and retail service directly provided to customers (http://www.wholefoodsmarket.com/blog/whole-story).

3) **Marriott On the Move**: Marriott on the move is by J.W. Bill Marriott, Jr., the Executive Chairman and Chairman of the Board of Marriott International, Inc., one of the world's largest lodging companies. It provides an international standard of service, accommodation and facilities (http://www.blogs.marriott.com/).
4) **The Cleanest Line**: This is a blog for the employees, friends and customers of the outdoor clothing company Patagonia. This company manages the research and development, design, manufacturing, merchandising and sales of outdoor clothing (http://www.thecleanestline.com/).
5) **Toyota Open Road Blog**: This the blog for Toyota Motor Corporation, a Japanese automotive manufacturer headquartered in Japan. They also perform all manufacturer scheduled maintenance, out of warranty repairs, as well as offer competitive prices on tires, batteries, etc. (http://blog.toyota.com retrieved in 2012).
6) **Red Cross Chat**: Red Cross Chat is the official American Red Cross Blog, with a goal of offering people valuable information about preventing, preparing for, and responding to emergencies (http://redcrosschat.org/).
7) **FedEx Citizenship Blog**: This is a blog for FedEx Corporation, an American company, which provides customers and businesses worldwide with a broad portfolio of transportation, e-commerce and business services (http://www.citizenshipblogs.fedex.designcdt.com retrieved in 2012).
8) **Facebook Blog**: This blog is for the Facebook, a social utility that provides an online social networking service to connect people to each other. It ties closely with an e-commerce platform, and provides avenues for advertising, branding, and marketing the products (http://www.facebook-blog.com/).
9) **Turkey Hill**: This is a blog for Turkey Hill Dairy, an American brand making frozen dairy products including ice creams, frozen yogurt, iced teas and fruit drinks distributed throughout the United States. The company sells its products in more than 49 states around the world through distributors and dealers (http://icecreamjournal.turkeyhill.com/).

10) **Lede**: This is a blog for the New York Times (NYT), an American media company which publishes newspapers. It belongs to the publishing industry which consists of companies engaged in the publishing of books, magazines, newspapers, directories, maps, etc. (http://thelede.blogs.nytimes.com/ retrieved in 2012).
11) **Wallet**: This is a blog for the Wall Street Journal, a daily newspaper with a special emphasis on business and economic news published in New York (http://blogs.wsj.com/wallet/ retrieved in 2012).
12) **Comment is free | guardian.co.uk**: This is a blog for the Guardian, a British national daily newspaper (http://www.theguardian.com/uk/commentisfree).
13) **Amazon Web Services Blog**: This is a blog for Amazon.com, Inc., an American e-commerce company, starting as an online bookstore, and now being the largest Internet-based retailer in the United States (http://aws.typepad.com/aws/ retrieved in 2012).
14) **Avaya - The Blog**: This is a blog for Avaya, a leading global provider of technologies for unified communications and collaboration, contact centre and customer experience management, and networking, along with related services to enterprises and government organizations around the world (http://www.avaya.com/blogs/).
15) **Brand Channeler**: Brand Channeler is a leading Brand Marketing blog from Cone Communications, a public relations and marketing agency offering services in the industry within a handful of areas of expertise (http://www.conecomm.com/tom-schimoler/).
16) **BreakingPoint Labs Blog**: This is a blog for Ixia, a company providing application performance (successfully deploying applications across network) and security resilience (ensuring networks are durable against ever-growing IT security threats) solutions and services to the organizations (http://www.breakingpointsystem.com/community/blog retrieved in 2012).
17) **Business Property**: This blog is for a company from the real estate industry. (http://www.hansmetal.com retrieved in 2012)
18) **Clorox-Dr. Laundry:** This is a blog for the Clorox company, which is a leading manufacturer and marketer of bleach and cleaning products (http://www.drlaundryblog.com).

19) **ComScore Voice**: This is a blog for ComScore, Inc., an American Internet analytics company providing such information as marketing data and analytics to many large enterprises. (http://www.comscore.com/Insights/Blog).

20) **Converge-A Bader Rutter Blog**: This blog is for Bader Rutter, which is a marketing communications agency that provides information services and helps brands thrive to companies from the insurance, construction, food and other B2B industries (http://www.bader-rutter.com/blog/).

21) **Corporate eye**: This blog is for the corporate eye company, which provides a number of services to the companies, e.g., corporate communications, marketing communications (B2B) and reputation & risk management. These services are able to enhance the practice standards of companies' communications (http://www.corporate-eye.com/main/blog/).

22) **Dr. Laskys Blog**: This blog is for the Indium Corporation which is a materials manufacturer and supplier to the global electronics, semiconductor, solar, thin-film and thermal management markets (http://blogs.indium.com/blog/Dr-Lasky-Blog/).

23) **Fiskars Ambassador**: This is a blog for Fiskars, a company from Finland bringing tools, such as knives, and utensils, that empower people in the garden, home, office, and classroom. It fills a lot of different needs, both manufacturer and consumer based. (http://www.fiskateers.com/)

24) **Fandango Movie Blog**: This is a blog for an online movie website providing movie theater showtimes, movie reviews and services of buying movie tickets in advance. It is a movie ticket retailer, which links consumers to online e-commerce businesses (http://www.fandango.com/fandangoblog retrieved in 2012).

25) **Garmin Blog**: This blog provides product news, events and behind-the-scenes information from Garmin Company, a leading manufacturer in GPS products (http://garmin.blogs.com/my_weblog/).

26) **Hu Yoshida**: This blog is for HITACHI, a Japanese company which is a globally recognized original equipment manufacturer. It offers a series of automotive aftermarket products and systems including electric power train systems, drive control systems, engine management systems, rotating electrical and car information systems. It is basically associated with manufacturing which involves the transformation of materials into

new products (http://blogs.hds.com/hu/).

27) **IAGblog**: This blog is for IAG, a company providing three primary services: Analytics, Consulting and Emerging Media. The activities conducted by this company is primarily supplying information services (http://www.iagblog.com/).

28) **Infusionsoft Blog:** Infusionsoft's blog delivers creative, innovative, and helpful ideas for small business. Offering great ideas is also a kind of information service activity (http://www.infusionblog.com/).

29) **JNJ BTW**: This is the blog for Johnson & Johnson, a healthcare company providing medical products and services. A broad range of products are provided by this company, such as those used in the baby care, skin care, oral care, and some nutritional and pharmaceutical products. Although some after-sale services are also provided, it is primarily a manufacturing company having actual productions (http://www.blogjnj.com/).

30) **Logos Bible Software Blog**: This blog is for an online website which offers the world's premier Bible study software, along with an entire line of resources for Christian living. Its main tasks include the offering of software, services and some trainings (https://blog.logos.com/).

31) **Marketing Conversation**: This blog is for the Gerris digital, a full-service digital strategy firm which offers its clients comprehensive online conversation marketing campaigns. It is primarily concerned with information and communication technologies (http://www.marketconversation.com/).

32) **Oracle Blogs:** This blog is for the Oracle Corporation, an American computer technology corporation providing a comprehensive and fully integrated pack of cloud applications, platform services, and software & hardware systems. Providing such internet service and software is primarily an information technology service (http://www.oracle.com/us/corporate/blogs/index.html).

33) **Rubbermaid - Adventures in Organization:** This blog is for Rubbermaid, an American Manufacturer and distributor of some kitchen, hardware, seasonal, and cleaning products, as well as commercial items (http://www.rubbermaid.com/en-US).

34) **RS Consulting blog**: This blog is for RS Consulting, a market research agency based in London, UK. This company helps individuals and businesses to acquire their information

and communication needs (http://www.rsconsulting.com/blog/).

35) **PurinaCare Pet Insurance Blog**: This blog is for PurinaCare pet health insurance, which provides pets insurance plans to customers. It is a leader in the pet products industry, being dedicated to improving the lives of cats and dogs through quality nutrition and care. This insurance company primarily focuses on providing services (http://www.purinacareblog.com/ retrieved in 2012).

36) **The LinkedIn Blog**: This blog is for the LinkedIn, which is an online social network for business professionals. It is designed specifically for professional networking - finding a job, discovering sales leads, connecting with potential business partners - rather than simply making friends or sharing media like videos or music (http://blog.linkedin.com/).

37) **Tyson Foods Hunger Relief**: This blog is for Tyson Foods, Inc., one of the world's largest processors and marketers of chicken, beef and pork, as well as prepared foods. This company is basically concerned with the manufacture of food products (http://www.tysonhungerrelief.com/blog/).

38) **The Unboxed Thoughts Blog**: This is the official blog of Prosek Partners, a public relations and financial communications consultancy serving a group of regional, national and international clients. It works with companies on their strategies to communicate with shareholders and deal with crises. (http://www.unboxedthoughts.com/)

39) **Upside Learning Blog**: This blog is for Upside learning, a learning technology solutions company, which help organizations to improve performance through a wide range of learning products and services. This company focuses on information communication, aiming to bring the experience and expertise to the organizations and help them succeed. (http://www.upsidelearning.com/blog/)

40) **Verizon-PolicyBlog**: This blog is for Verizon Communications Inc., an American broadband and telecommunications company. The main tasks include designing, building and operating global networks, information systems and mobile technologies that connect people, grow businesses and economies, and improve communities. It has the largest market share in the wireless telecommunications industry. (http://policyblog.verizon.com/).

41) **Xerox Blogs**: This blog is for the Xerox Corporation, the world's leading enterprise providing not only business services, but also digital printing equipment. The backbone of

this company, including many other companies, is its technology; with a focus on communication and customer service, the information technology professionals of Xerox Corporation keep the organization running (http://www.blogs.xerox.com/).

4.3 Corpus design

The creation of CBC has followed some guiding principles from the literature that are not only theoretically but also practically related. More systematically, the guiding principles adopted for the corpus building in this research project constitute a model, in which a coherent sequence of phases covering the 'what' and the 'how' of corpus building is included. I followed two major stages in corpus building: corpus design and corpus creation (Sinclair, 2005). Like any other corpora, the design of CBC is for particular purposes; the usefulness of it must be judged with regard to its purposes, and planned in accordance with some crucial aspects in the guiding principles, including corpus representativeness, corpus size, balance and some ethical issues, etc. (e.g., Biber, 1993; Biber et al., 1998; Cheng, 2011; Leech, 1991; Nelson, 1996). In this way, the theoretical underpinning for the decisions made in the particular designing stages could possibly give rise to a more sensible implementation stage.

The starting point is the identification of the goal of the creation of CBC. This is a fundamental step in corpus design which may influence all subsequent decisions. CBC is among the thousands of corpora in the world, being created for specific research projects. In the current research, the corpus-based analysis aims to contribute to the register analysis of corporate blogs. As a linguistic dimension, register is "a contextual category correlating groupings of linguistic features with recurrent situational features" (Gregory & Carroll, 1978, p. 4) and interacts with genre since texts can be viewed as the product of discourse community. As indicated by the research questions raised in this research project, the register analysis of corporate blogs focuses on the lexical and the grammatical level of language. And the corpus-linguistic method is an applicable approach to this research, in that it allows for the study of a large number of linguistic features in a "collection of authentic, naturally occurring language texts" (Sinclair, 1991, p. 171). This method is supported by adequate tools, as well as a good corpus design which serves to answer research questions. In the course of the

corpus design of CBC, I accepted an idea that a good corpus should 'represent' the language variety (Sinclair, 2005), although it is difficult to judge whether or not a corpus is better than another. The 'representativeness' of a corpus is closely related with such issues like corpus size (Biber, 1993), balance (Hunston, 2002; Kennedy, 1998), diversity (Biber et al., 1998), and some ethical considerations. A representative corpus must represent a language with varied linguistic distributions and encompassing various text types, while a balanced corpus should attempt to capture their different percentage levels. This is somewhat difficult to achieve, mainly because it is difficult to precisely know all the text types and their proportions of use in the corpus (McEnery & Wilson, 2001). However, in light of the issues of corpus building in previous studies, it is assumed that the corpus can be representative of any genres or subject fields, although there are possibly some practical constraints, such as a the unavailability of machine-readable text, or copyright restrictions, and it is not always possible to assemble the representative and balanced corpus ideally wanted.

In order to construct the CBC, the first issue I took into account is the overall size of the corpus, which is inseparable from the corpus type it could be categorized into. CBC is a self-built rather than an existing corpus, and it could be categorized into an existing corpus type, i.e., the specialized corpus, because it represents a particular variety of language from a specific domain (e.g. business, biology, social sciences) and used for a particular purpose. Similar examples include academic weblogs (Stuart, 2006), political blogs (Aleknavičiūtė, 2009), etc. Accordingly, a specialized corpus differs from a general corpus to the extent how the knowledge is shared. They also differ from each other in their size; a general corpus is often in a large scale, such as the 100-million-word British National Corpus, the Brown Corpus of Standard American English, while a specialized corpus is usually much smaller than that, because it represents smaller subsections of language than general corpora (Hunston, 2002). There are not any strict rules showing what might be an appropriate size for a specialized corpus. In spite of this, some principles from the literature were followed to deal with the issue of the size in corpus design. The trend from an early prevailing view of "bigger means better" (Leech, 1991, p. 9) to the emphasis on the use of a smaller corpus (e.g., Ghadessy, Henry, & Roseberry, 2001; Hunston, 2002) suggests a plausible direction for the corpus design of CBC.

It has been suggested that “there is no point in having bigger and bigger corpora if you cannot work with the output” (Kennedy, 1998, p. 68), as illustrated in some empirical studies. For example, BEC (business English corpus) consists of 20,000 words, the size of which is considered to be appropriate enough to provide statistically reliable results (Nelson, 2006). Besides, the size of some specialized corpora in blog research (e.g., Grieve et al., 2011; Nowson et al., 2005; Stuart, 2006) varies from 410,000 running words (Nowson et al., 2005) to 16 million words (Stuart, 2006). As for CBC, which is a specialized corpus, it is more difficult and time consuming to obtain blog texts as opposed to general texts. It may even reach “closure” when an increase in the size of this corpus does not bring in new instances of the given feature (McEnery & Wilson, 2001, p. 166). Taken together, the size of CBC which has 590,520 words is assumed to be sufficiently large.

Apart from the overall size of the corpus, the component of the corpus is another issue that I took into account in corpus building. In particular, the technique of ‘sampling frames’ (Biber, 1993) was taken into account, which includes several kinds of probabilistic samples: simple random sampling and stratified sampling. What has been adopted in the design of CBC is neither a “simple random sampling”, in which “all sampling units within the sampling frame are numbered and the sample is chosen by use of a table of random numbers”, nor a “stratified random sampling”, by means of which the sample is “divided into relatively homogeneous groups (so-called strata) and then samples each stratum at random (Dale, Moisl, & Somers, 2000, p. 152). The design of CBC potentially uses both approaches, being partially attributed to its nature as a web data. The corporate blogs are not accessed from those public search engines, such as Google, Yahoo, but collected from BlogRank (http://www.blogmetrics.org/Corporate_Blogs), a website in which the top-ranked corporate blogs are listed in one place. In BlogRank, the top-ranked corporate blogs are categorized and listed on the basis of more than 20 different factors, such as incoming links, Technorati ranking, unique monthly visitors, ultimate ranking, etc. What has been chosen for the present research are the 50 top-ranked corporate blogs based on the ultimate ranking. In each corporate blog, the posts are listed in accordance with a reversed chronological sequence; thus, the posts selected from each corporate blog are naturally divided into relatively homogeneous groups. In

each corporate blog, the selection of posts has relied more heavily on a “simple random sampling” approach, by means of which those sequenced posts are chosen one after another. This selection mainly depends on the extra-linguistic feature of time.

Another issue associated with corpus size is whether the whole texts or the fragments from preselected texts should be included. I decided to include the whole texts rather than the fragments of CB texts, on account of the advantages and disadvantages of the inclusion of whole texts. On the one hand, including the whole texts is certainly useful in discourse or genre studies because certain linguistic characteristics are typical of certain parts of a text. On the other hand, the disadvantage is this may challenge some copyright issues, and give rise to an imbalance of the corpus. Sinclair (1991), a strong advocate for whole texts in corpora, has argued:

> The alternative is to gather whole documents. Then there is no worry about the marked differences that have been noted between different parts of a text. Not many features of a book-length text are diffused evenly throughout, and a corpus made up of whole documents is open to a wider range of linguistic studies than a short collection of samples. (p. 19)

In terms of the design of CBC, I also advocate the inclusion of whole texts. I consider that very naturally occurring texts with varied length, rather than those texts with the same number of words should be retrieved and retained for corpus building. In practice, the stability of distributions of linguistic features, and the text fragments with the same number of words (e.g., 2000 running words) that may be appropriate for the study (Biber, 1993, p. 252) are not easy to be achieved, since the length of CB texts does vary naturally and greatly. The mean word length of CB texts is 507 words, while the shortest and longest texts contain 103 and 2710 words respectively. These data indicate a great variation of the number of words in CB texts, which is its natural attribute. Another reason why this is applicable is that the permission of the use of full texts is granted by copyright holders. The corporate blog is among the publicly available sources for a corpus study. As for the ethical issues of publicly available sources, e.g., online blogs, some problems and controversies may still exist (Bowker & Tuffin, 2004; King, 1996; Sixsmith & Murray, 2001). An important aspect to the ethical problems of online research

is what the notions of private and public mean in online research, which has no consensus among social scientists (Hutchinson, 2001). Some researchers argue that online material is publicly available and, thus falling in public space and participant consent is not necessary (Sudweeks & Rafaeli, 1996). Other researchers claim that online content should be treated as a 'private' form of interaction, and require consent from participants (Elgesem, 2002; King, 1996). There are also some researchers who argue that there is no clear distinction between public and private online interactions (Waskul, 1996). As for corporate blogs, it is located in the public; however, it is not necessary to obtain consent, because the corporate blog is for a research rather than a commercial purpose in the current research, which means authors are not advantaged in any way, and there is no financial gain involved (Hookway, 2008). Accordingly, the whole texts from corporate blogs can be retrieved legally.

Some other issues closely associated with the representativeness of corpus creation are also taken into account. They are balance (Hunston, 2002; Kennedy, 1998) which refers to the "weighting between the different sections" (Kennedy, 1998, p. 62) in a corpus and diversity (Biber et al., 1998), which is related with "the range of text types in a language, and the range of linguistic distributions in a language" (Biber, 1993, p. 243). Both of these criteria play an important role in helping the corpus represent the language or a language variety. To ensure the balance of CBC, the following text selection criteria were used: "domain", "time" and "medium" (Aston & Burnard, 1998, pp. 29-30). Domain refers to the content type (i.e. subject field) of the text; time refers to the period of text production, while medium refers to the type of text publication such as books, periodicals or unpublished manuscripts (Aston & Burnard, 1998; McEnery, Xiao, & Tono, 2006). In a general sense, domains of written texts are broad 'subject fields' (Burnard, 1995), covering broader context categories, particularly the major spheres of social life (leisure, business, education, and institutional/public contexts). For a specialized corpus like CBC, which is from the domain of business, the focus can be put on the range of subject topics it covered. It is assumed that CBC involves a wide range of topics since the 41 corporate blogs in CBC are from 41 enterprises that could be categorized into many different industries, such as high tech industry, service industry, sales industry, manufacturing industry. All these initial observations can be justified in the later classification on the basis of

some global industry classification systems (cf. Section 5.6). CB texts from different enterprises have different products, services and information to provide. Even in a corporate blog, some posts are delivered for sharing some advice and insights into their way of doing things, some posts bring customers into their personal stories, and some posts announce new products and services. Overall, corporate blogs in CBC have varied aims of text production. Secondly, the period of text production in CBC lasts for approximately two years from Aug. 2009 to May 2012, while the collection period started from May 2012, and ended in August 2012. Thirdly, as an online media, the corporate blog includes increasing sources of language and corpus linguistics data, which distinguishes it from traditional type of text publication. The balance of a web-based corpus, including CBC, can also be assessed in accordance with the factor of the single or multiple authorship (Sharoff, 2006). Among the 1020 CB texts for the current research, 145 of them are co-authored, and the vast majority of them are single authored. This proportion, I suppose, does not have much distortion to the balance of the corpus, since the criterion of the range of “domain”, “time” and “medium” (Aston & Burnard, 1998, pp. 29-30) plays a more important role for corpus building, and meeting these criteria simultaneously cannot be achieved. Besides, this authorship is acceptable in that the line between the single-author and the co-authors would be blurred when the single-author and the co-authors represent the company behind them.

Besides, to ensure diversity in a corpus, the two areas suggested by Biber (1998) can be considered: register variation and subject matter. The issue of subject matter somewhat overlaps with the text selection criterion of “domain” in Aston and Burnard (1998, pp. 29-30), while the use of different registers depends on different communicative situations. To some extent, register is more concerned with language-internal criteria (situational and linguistic variability), the essential elements for a corpus to be representative (Biber, 1993). In the current research, it is difficult to judge how many registers have been included in CBC because the data are from various websites, instead of some pre-categorized registers. In spite of this, it is still assumed that CBC is composed of different registers, because the communicative situations in corporate blogs are varied. As what has been discussed in the issue of balance, communication connects an enterprise to their customers, in which different

types of communication serve different purpose, e.g., reinforcing the company brand, providing customer services, or increasing sales of the company's products. While the real situation is it is impossible to create a perfectly representative corpus (Hunston, 2002; Kennedy, 1998; Nelson & West, 2000), or more specifically, ensuring representativeness in obtaining data is not an easy task, CBC has been designed as representative as possible.

To sum up, guided by these theoretically and practically related principles, CBC was created. And presented below is the corpus building procedure of CBC.

1. Access a blog entry, which usually includes any combination of text, images, links, comments etc…, as shown in Figure 4.3.

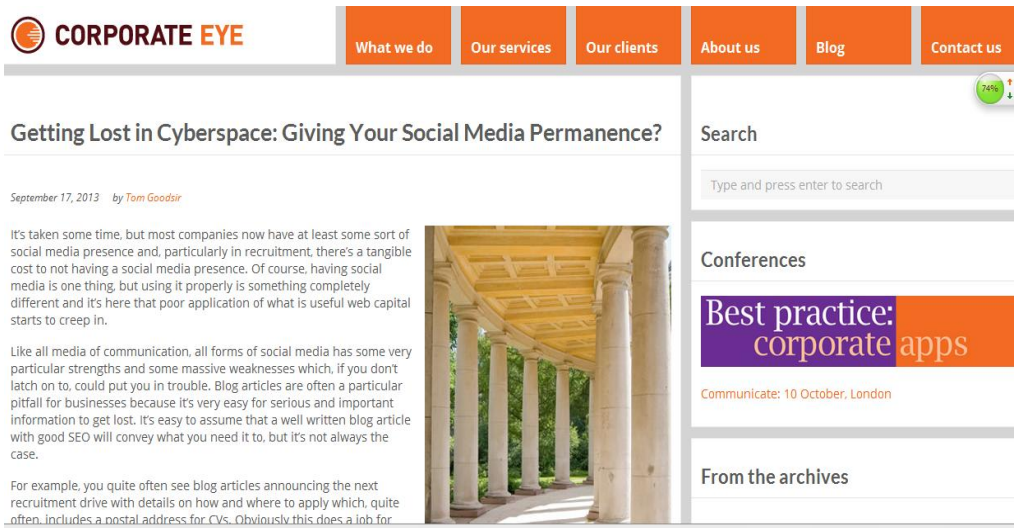


Figure 4.3: The corporate blog from Corporate Eye

This step is easy to achieve, because the 50 top-ranked corporate blogs used for corpus building are from BlogRank (http://www.blogmetrics.org/Corporate_Blogs). If clicking on any blog titles, it takes us to an external resource, i.e., the homepage of a corporate blog (see Figure 4.3). Before retrieving raw data from the internet, it is decided that 41 out of 50 top ranking corporate blogs should be retained as sources for data collection because the other 9 were either written in non-English languages, or dominated by non-textual content, and even non-accessible.

2. Retrieve raw data from internet

In order to retrieve raw data from a corporate blog, I firstly retained the textual content from corporate blogs which usually include a variety of content, i.e., textual, visual, audio, and commentary part; and then, in order to get pure body texts needed for the present study, I

removed several components other than the pure texts in a post, i.e. a short, descriptive title, the date and time it was posted, or some comments by the blog readers. Totally, there are 1058 texts selected to represent corporate blog. They are the most recent CB texts posted online collected in accordance with the timeline, specifically, from Aug. 2009 to May 2012. The CB texts are also kept at their original length, on the basis of a consideration that whole texts are able to provide more opportunities for words, longer texts, as well as variety of vocabulary to occur more frequently (Coxhead, 1998). On this basis, 38 CB texts were removed from the original 1058 texts, because the word count in the 38 texts is less than 100. Finally, a CBC consisting of 1020 pure CB texts is created. The texts collected in CBC show a varied word count from 103 to 2710, and the average word count is 507.

3. Create directory files

After the creation of CBC, what comes next is to create 41 directory files, each of which includes blog posts from the same website and represents a database corresponding to a corporate blog. Inside each directory, a set of txt files are contained, which corresponds to a number of pure CB texts.

4. Arrange text data in table

When database were prepared, text data is able to be arranged in the form of excel tables as a record. A series of information are classified in accordance with such criteria as title, postdate, author's name, gender of the author, company name, the industry of the company, date of data collection, links to txt text files, etc. In the following table, some details of CB texts from two corporate blogs are presented.

Table 4.2: CB texts from two corporate blogs

Title of CB texts	Publishing time	Author's name	Gender	Organization
Say Hello To Mobile Baggage Tracking	2011/11/22	Jonathan Mayfield	male	Delta
New Boarding Pass Debuts On Delta.com & Airport Kiosks	2012/2/1	Rick E.	male	Delta
Wi-Fi Update – The Internet Flies Delta Connection!	2011/9/7	Chris B.	female	Delta
Groupon A Delta Sky Club Pass & Go!	2011/10/25	Joy H.	male	Delta
More Economy Comfort Means More Room to Relax!	2011/10/20	Chris B.	female	Delta

The New Delta BlackBerry App is Here!	2011/10/4	Jonathan Mayfield	male	Delta
Delta's 767-300ER Fleet Getting a New Look	2011/8/5	Chris B.	female	Delta
View from 30,000 Feet: Delta on Diversity	2011/6/23	Trebor Banstetter	male	Delta
Military Travel: Baggage Policies & Our Thoughts	2011/6/7	Rachael R	male	Delta
Economy Comfort Installation Updates	2011/5/5	Chris B	male	Delta
Relax, Stretch Out & Get Comfortable	2011/2/7	Chris B	male	Delta
How To: Rebook Your Travel Plans	2011/1/7	Jerry F	male	Delta
Boeing 747 Enhancements: Flat Beds, Touch Screen TVs and More	2010/9/2	Mike B	male	Delta
We're Listening. And Delivering.	2010/7/29	Jeff Robertson	male	Delta
25 Years of MaDDogs	2012/4/30	Laura M	female	Delta
Your Trip (Extras) – Your Choice	2012/3/13	Cheryl B	female	Delta
Behind-the-Wing: Passenger Stories	2012/4/19	Laura M	female	Delta
Behind-the-wings: The Pilot Life At Delta	2012/3/23	Dawn C	female	Delta
A New Year Means More Ways Than Ever to Surf the Internet on Delta!	2012/1/6	Chris B	female	Delta
Fare Displays On delta.com	2012/5/18	Bob Kupbens	male	Delta
Cooking With Almond Butter	2012/2/20	Alana Sugar	female	Whole Foods Market
Just Beet It	2012/2/16	Mary Olivar	female	Whole Foods Market
Help a Million More People Change Their Lives	2012/2/22	Whole Planet Foundation	co-authors	Whole Foods Market
Eating Near the Poverty Line…at Whole Foods Market	2012/2/22	Lisa Johnson	female	Whole Foods Market
Sweet Recipes for Sweet Potatoes	2012/2/13	Alana Sugar	female	Whole Foods Market
Putumayo Celebrates Carnival in Rio with Brazilian Beat	2012/2/11	Dan Storper	male	Whole Foods Market
What's New in the Grocery Aisles: February	2012/2/8	Elizabeth Smith	female	Whole Foods Market

Natural Ingredients for Energy	2010/3/16	Chris Jensen	male	Whole Foods Market
GMO/Monsanto Buyout Rumors Untrue	2012/2/2	Libba Letton	female	Whole Foods Market
Your Favorite Holiday Food Tradition	2011/12/12	Elizabeth Beal	female	Whole Foods Market
Win Satsumas for Sharing	2011/12/5	Jennifer Cheng	female	Whole Foods Market
Enjoy A Celtic Christmas from Putumayo	2011/12/3	Dan Storper	male	Whole Foods Market
The Thanksgiving I'm Most Thankful For Contest	2011/11/14	Anna Madrona	female	Whole Foods Market
Seasonal Wellness Shopping List	2011/10/29	Chris Jensen	male	Whole Foods Market
Let Us Stock Your Thanksgiving Pantry	201110/28	Allison Burch	female	Whole Foods Market
Homeopathy for Cold and Flu Season	2011/10/18	Chris Jensen	male	Whole Foods Market
Putumayo's Acoustic Café	2011/10/8	Dan Storper	male	Whole Foods Market
$5 Million in Loans to Local Producers	2011/10/6	Heather Kennedy	female	Whole Foods Market
Organic Valley	2011/9/20	Marc Hamel and Ha Lam	team work	Whole Foods Market
The Annual "Best Ever" Guacamole Search	2012/1/30	James Parker	male	Whole Foods Market

In this table, it is shown that CB texts selected were listed in reverse chronological order. A further examination of all data revealed that the CB texts collected in each corporate blog for the present project were not restricted to the same period, and not very evenly distributed. This is partly owing to their different updating frequency, and partly due to the availability of CB texts.

4.4 Data processing and statistical analysis

After the corpus was appropriately designed, it was then analysed by using some statistical approaches. The analysis of the corpus for this research project comprises of two main parts: a multidimensional analysis, and a statistical analysis of some lexical and grammatical features related to the metafunctions in SFL. The first part of the analysis aims to identify the

principal dimensions of linguistic variation and the primary text types in corporate blogs, together with functional associations based on the relationship between CB text types and industries. The procedures of this analysis are described in Section 4.4.1. The second part of the analysis aims to identify how lexical and grammatical features distinguish different text types of corporate blog, and how they serve to realize the meanings in the context of situation. The procedures of this analysis are presented in Section 4.4.2.

4.4.1 Multi-dimensional analysis

In practice, several methodological steps were followed in MD analysis (e.g., Biber, 1995; Ludeling & Kyto, 2008), as summarized in the following sections.

4.4.1.1 Selection of linguistic features

Selecting linguistic features is one of the crucial steps in multi-dimensional analysis. Biber's multidimensional analysis requires the principled selection of a large and diversified number of language features. For example, the frequency and distribution of 67 lexical and syntactic features were measured in 481 spoken and written texts in Biber's (1988) study. In the current MD analysis, the CB texts are from a more specialized domain, and thus it possibly includes a limited range of variation than that in Biber's (1988) studies of written and spoken British English. In spite of this, the linguistic features included in this analysis should also be diversified, and typically associated with some form-function associations.

In this MD analysis, 32 linguistic features from Biber's initial 67 language features were selected, as discussed in Chapter 3 and listed in Appendix 2. The choice of these linguistic features for the current analysis was partly determined by the availability for tagging these features by using some reliable computer programs, and partly based on the following considerations.

Firstly, the selection of 32 linguistic features has followed a basic criterion proposed by Gorsuch (1988, p. 237), i.e., "marker variables". Marker variables are those variables that better represent the factor, conveying the meaning of the underlying dimension without being

related to any other factors (Cattell, 1978; Gorsuch, 1983). This definition shows two distinct fundamental properties of marker variables: representativeness and simplicity, which means the marker variables should be highly related to the factor, and each variable is related only to one factor. The variables included in a factor analysis should also be correlated and reliable, and thus, accepted marker variables can greatly enhance the confidence in the set of variables (Nunnally & Bernstein, 1978).

Following these principles, 32 linguistic features were selected from Biber's (e.g., 1988, 2006) work which is supported by a large amount of background studies, particularly the previous comparisons of spoken and written texts (e.g., Chafe & Tannen, 1987), and descriptive grammars of English (Quirk et al., 1985). All the linguistic features included in the current MD analysis are marker variables in that they are highly related to the factors, such as present tense verbs (factor loading 0.86), second-person pronouns (factor loading 0.86), demonstrative pronouns (factor loading 0.76), first-person pronouns (factor loading 0.74), pronoun it (factor loading 0.71), indefinite pronouns (factor loading 0.62), possibility modals (factor loading 0.50) from Dimension 1, past tense verbs (factor loading 0.90), third-person pronouns (factor loading 0.73), perfect aspect verbs (factor loading 0.48) from Dimension 2, time adverbials (factor loading 0.60), and place adverbials (factor loading 0.49) from Dimension 3 (Biber, 1988). Taken together, the language features selected have been proved to be significant statistically in Biber's (1988) initial studies on spoken and written language in English, with relatively higher loadings in the six basic dimensions/factors of variation in English language, i.e. taking up a higher percentage of variance in that the factor/dimension. The inclusion of these linguistic features makes possible a statistically significant MD analysis in the current research.

Besides, some correlated variables that have been explicitly viewed as indicators for the interpretation of certain text types in some earlier work by Biber (e.g. 1988, 2006) are also selected and then adopted for the current research. For example,

- Argumentative texts are characterized by the use of *modals*, e.g., prediction, necessity, and possibility modals, which is "designed to persuade the addressee" (Biber, 1988, p.

111).

- *Past tense* is a linguistic feature that serves as an indicator distinguishing narrative and expository texts that are basically “active and event-oriented”, and “more static and descriptive” respectively (Biber, 1988, p. 109).
- The features *high frequency of nouns, high type/token ratio, and the presence of nominalizations and passives* are further indicators of expository texts since they are characterized by a highly informational discourse (Biber, 1988, p. 104).
- Texts grouped into narrative texts have a high frequency of *past tense forms, third-person personal pronouns, and perfect aspect verbs* (Biber, 2006)

Secondly, the linguistic features selected for the present study are from a wide range of 12 grammatical/functional categories[9], thus being considered as potentially important for this multi-dimensional analysis. Overall, the inclusion of this group of features is based on a review of previous theoretical and empirical works on register variation and related lines of inquiry such as study of the differences of written and spoken communication.

Thirdly, another design decision which concerns the selection of sample size in a factor analysis were also adopted. In the literature, a wide range of recommendations are proposed as for either the minimum necessary sample size, *N,* or the minimum ratio of *N* to the number of variables being analysed, *p* (MacCallum et al., 1999). There is a dramatic variation in these recommendations which causes them to be of rather limited value to empirical studies. For example, Gorsuch (1983) and Kline (1979) recommended that *N* should be at least 100, Guilford (1954) argued that *N* should be at least 200, and Cattell (1978) claimed that the minimum desirable *N* to be 250. In view of this, I did not take into account the issue of the selection of sample size in the present study. Rather, I focused on the number of variables when selecting linguistic features for this MD analysis. In order to include an ‘appropriate’ number of variables, the degree of over-determination of the common factors was measured, which refers to the degree to which each factor is clearly represented by a sufficient number of variables (MacCallum et al., 1999). It is a complex issue, as well as a crucial respect for the

[9]A. Tense and aspect markers B. Place and time adverbials C. Pronouns and pro-verbs D. Questions E. Nominal forms F. Passives G. Stative forms H. Subordination features J. Prepositional phrases, adjectives, and adverbs K. Lexical specificity L. Models

evaluation of the quality of a factor analysis. Previous studies focusing on the issue of over-determination have proved and suggested that the variables included cannot be too much without any limitations. I decided to follow this viewpoint and include a moderate amount of variables in the analysis. Some specific claims and suggests are presented below:

- Comrey and Lee (1992, pp. 206-209) discussed over-determination and recommended at least 5 times as many variables as factors.
- Without considering specific ratio of the number of variables to the number of factors, Tucker, Koopman, and Linn (1969) demonstrated that several aspects of factor analysis solutions can be intensively improved when the ratio of the number of variables to the number of factors increases. This effect has also been observed by Browne (1968) and by Velicer and Fava (1987, 1998).
- Either 6 or 7 variables per factor or a rather small number of factors is considered as high over-determination of factors (MacCallum et al., 1999).
- At least 4 measured variables for each common factor and perhaps as many as 6 (Fabrigar et al., 1999)
- A factor with fewer than 3 items might be weak and unstable (Costello & Osborne, 2011)

In the present MD analysis, 32 variables (linguistic features) is either a very large or a small number, and its 'appropriateness' can be tested after factor analysis is done, in which how many factors retained can be decided. In a general sense, the number of variables selected in each MD analysis is varied from each other; thus, it makes no sense to distinguish right and wrong between different numbers of variables (here linguistic features) selected in different studies. Rather, it is more plausible to follow some commonly agreed and basic criteria, e.g., reliabilities and expected correlations with other variables, to consider the feasibility of variable selection in practice, to select a moderate amount of variables, and finally to make sure the variables are potentially important for the factor analysis.

4.4.1.2 Text tagging and frequency counts of linguistic features

This section aims to give a more elaborated account of step 2 and step 3 in the MD analysis (Section 2.3.4.1). In step 2 (Computer programs are used to identify or 'tag' all relevant

linguistic features in the texts, and all texts are edited to ensure that the linguistic features can be accurately identified), 32 linguistic features in each CB text was automatically 'tagged' by using CLAWS 7, a part of speech (POS) tagging programme. It is a free POS tagging software for English text available on http://ucrel.lancs.ac.uk/claws/trial.html. It has been commented on the website http://ucrel.lancs.ac.uk/claws/ that CLAWS has consistently achieved 96-97% accuracy (the precise degree of accuracy varying according to the type of text). Besides, Marshall (1987) has also reported that the performance of CLAWS tagger is 94%. Approximately 65% of the words were tagged unambiguously, and the disambiguation program achieved better than 80% success on the ambiguous words. Before all texts were tagged, a pre-editing process was followed to ensure that the linguistic features are grammatically accurate, in case there are some special features of the CBC, such as non-standard spelling, or invented words. Tagging the various linguistic features to be used in factor analysis is a major component in MD studies (e.g., Biber, 1988). For the MD study on corporate blogs, there are enough electronically available texts; so all texts for this research were able to be entered automatically in any case. Presented below is an example of tagged text:

> This_DD1 week_NNT1 ,_, May_NPM1 took_VVD us_PPIO2 up_RP on_II the_AT offer_NN1 and_CC visited_VVD the_AT Dairy_NN1 with_IW her_APPGE son_NN1 and_CC granddaughter_NN1 for_IF a_AT1 tour_NN1 ._.
>
> Before_CS we_PPIS2 ended_VVD our_APPGE conversation_NN1 with_IW May_NP1 ,_, we_PPIS2 asked_VVD her_PPHO1 to_TO reveal_VVI her_APPGE secret_NN1 to_II a_AT1 long_JJ ,_, healthy_JJ life_NN1 ,_, to_II which_DDQ she_PPHS1 replied_VVD simply_RR ,_, Think_VV0 positive_JJ and_CC eat_VV0 ice_NN1 cream_NN1 ._. Hard_RR to_TO argue_VVI with_IW that_DD1 !_! (from file 8 turkeyhill)

In step 4 (Additional computer programs are run to compute normed frequency counts of each linguistic feature in each text of the corpus), the programme named AntConc was used. AntConc (http://www.laurenceanthony.net/software/antconc/) is a free, simple, and easy-to-use corpus analysis software. It has been adopted as a primary corpus linguistic software for this study, because it allows for the analysis of word count, keywords (words

determined to occur in a much higher frequency in a particular corpus compared with a particular reference corpus that must also be specified by the user, e.g., the Brown word list), and collocations (the co-occurrence of certain terms). In order to count 32 linguistic features, the following tools were used in AntConc, through which the data is able to be processed:

1) File Tool

The very first procedure was to select one or more files for processing by using the 'Open File(s)' or 'Open Dir' options in the 'File Menu'. Files were then continually added to the system. The list of selected files is shown in the left frame of the main window, as can be seen in Figure 4.4:

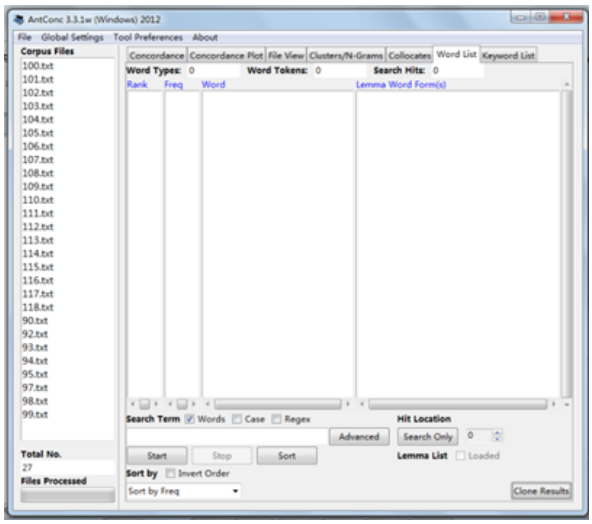


Figure 4.4: File tool

2) Word List / Keyword List Tools

By using the Word List / Keyword List Tools, I generated a list of all the words in the corpus. Hockey (2001) has claimed that an ideal word list generation program should be able to sort words into alphabetical or frequency order. The Word List Tool of AntConc offers these features and also the ability to count words based on their 'stem' forms. In order to produce a wordlist, I also chose a 'word list' option, and then pressed 'start' (Figure 4.5).

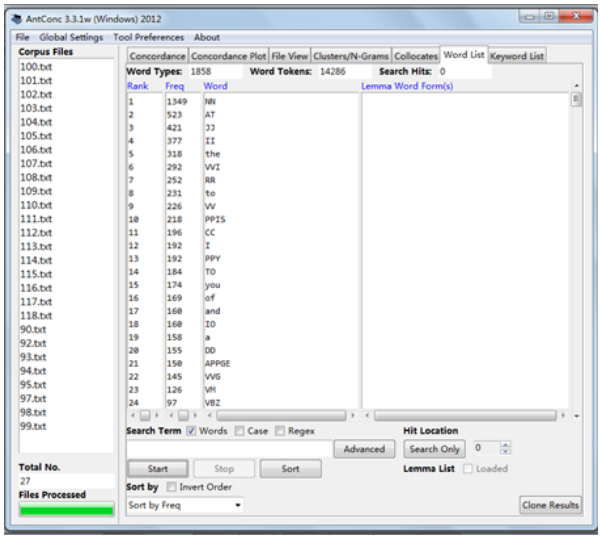


Figure 4.5: Wordlist tool

In order to count the linguistic features in each CB text more efficiently, I then chose 'choose by word', and pressed 'sort'. In this way, the words are listed in accordance with alphabet letter, as shown in Figure 4.6:

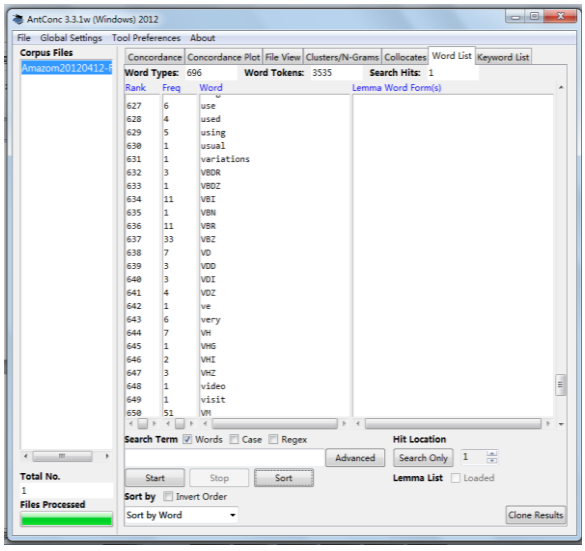


Figure 4.6: Sort tool

At this stage, a problem possibly occur, i.e., the frequency of some linguistic features cannot be identified directly and easily. For example, the frequency of the linguistic feature of *Past tense* from category A (Tense and aspect markers) cannot be easily judged and identified, because it may consist of several kinds of tag markers (VDD, VHD), or its co-occurrence with other verbs should be judged by the use of a concordance tool (**VBDR** + VVG e.g. were doing; **VBDZ** + VVG e.g. was doing), as shown in Figure 4.7:

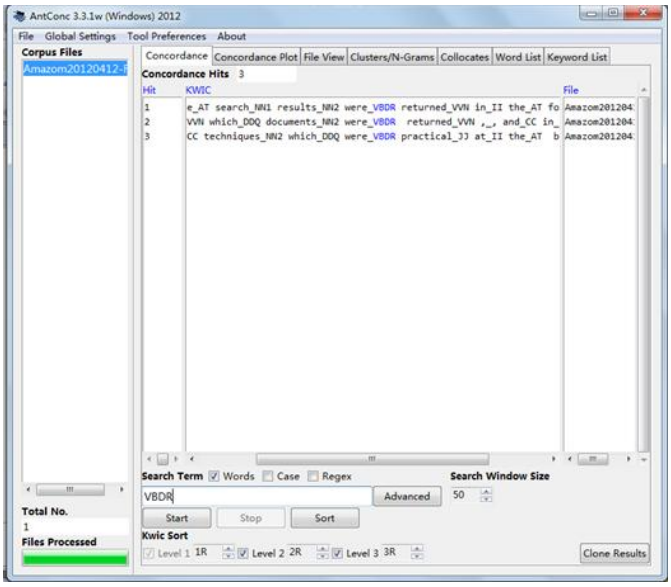


Figure 4.7: Concordance tool

It can be seen clearly that Figure 4.7 shows a screenshot of search results by using a 'concordance tool' in AntConc; there are three concordance lines consisting of VBDR, but none of them co-occur with VVN. With the help of the Concordance Tool, I was able to find a huge number of examples in varied contexts and situations quickly and efficiently. What is presented below is a table which can give us a full picture of the tagging for each linguistic feature used in the present MD analysis (excluding mean word length, and TTR, because they can be counted directly in AntConc and Microsoft Word).

Table 4.3: Tagging of linguistic features for the MD analysis

Linguistic features	Tagging	Definitions
1. Mean word length		
2. Type-token ratio (TTR)		
3. Past tense	**3.1 VBDR** + VVG (e.g. were doing)	were + ING/ -ing participle of lexical verb (e.g. giving, working)
	3.2 VBDZ + VVG (e.g. was doing)	was + ING -ing participle of lexical verb (e.g. giving, working)
	3.3 VDD	Did
	3.4 VHD	HAD， had (past tense)
4. Perfect aspect verbs	4.1 VH0 + VVN (e.g. have done)	HAVE， have, base form (finite) + past

		participle of lexical verb (e.g. given, worked)
	4.2 VHD + VVN (e.g. had done)	HAD，had (past tense) + past participle of lexical verb (e.g. given, worked)
	4.3 VH0 +VBN (e.g. have been)	HAVE，have, base form (finite) + been
	4.4 VHD+ VBN (e.g. had been)	HAD，had (past tense) + been
5. Present tense	5.1 VV0	base form of lexical verb (e.g. give, work)
	5.2 VBM + VVG (e.g. am doing)	am + ING -ing participle of lexical verb (e.g. giving, working)
6. Agentless passive	**6.1 VBDZ** + VVN（e.g. was told）/ **Agentless passive**	was + past tense of lexical verb (e.g. gave, worked)
	6.2 VBDR + VVN (e.g. were told)/ **Agentless passive**	were + past tense of lexical verb (e.g. gave, worked)
	6.3 VBM + VVN (e.g. am asked)/ **Agentless passive**	am + past tense of lexical verb (e.g. gave, worked)
	6.4 VH0 + VBN + VVN (e.g. have been told) / **Agentless passive**	HAVE，have, base form (finite) + been + past participle of lexical verb (e.g. given, worked)
	6.5 VHD + VBN + VVN (e.g. had been told) / **Agentless passive**	HAD，had (past participle) + been + past participle of lexical verb (e.g. given, worked)
7. By passives	**7.1 VBDZ** + VVN（e.g. was told）/ **+ by**	was + past tense of lexical verb (e.g. gave, worked)
	7.2 VBDR + VVN (e.g. were told)/ **+ by**	were + past tense of lexical verb (e.g. gave, worked)
	7.3 VBM + VVN	am + past tense of

	(e.g. am asked)/ **+ by**	lexical verb (e.g. gave, worked)
	7.4 VH0 + VBN + VVN (e.g. have been told) / **+ by**	HAVE，have, base form (finite) + been + past participle of lexical verb (e.g. given, worked)
	7.5 VHD + VBN + VVN (e.g. had been told) / **+ by**	HAD，had (past participle) + been + past participle of lexical verb (e.g. given, worked)
8. First-person pronouns	PPIO1	1st person sing. objective personal pronoun (me)
	PPIO2	1st person plural objective personal pronoun (us)
	PPIS1	1st person sing. subjective personal pronoun (I)
	PPIS2	1st person plural subjective personal pronoun (we)
9. Second-person pronouns	PPY	2nd person personal pronoun (you)
10.Third-person personal pronouns	PPHO2	3rd person plural objective personal pronoun (them)
	PPHS1	3rd person sing. subjective personal pronoun (he, she)
	PPHS2	3rd person plural subjective personal pronoun (they)
11. Pronoun *it*	PPH1	3rd person sing. neuter personal pronoun (it)
12. Demonstrative pronouns (that, this, these, those as pronouns)	DD (by using a cluster tool, the frequency of that, this, these, those as pronouns can be added)	singular determiner (e.g. this, that, another) plural determiner (these, those)
13．Indefinite pronouns (e.g.,	PN	indefinite pronoun,

anybody, nothing, someone)		neutral for number (none)
	PN1	indefinite pronoun, singular (e.g. anyone, everything, nobody, one)
14. Total prepositional phrases	IFF	for (as preposition)
	II	general preposition
	IO	of (as preposition)
	IW	with, without (as prepositions)
15.Place adverbials (e.g., above, beside, outdoors)	RL	locative adverb (e.g. alongside, forward)
16. Time adverbials (e.g., early, instantly, soon)	RT	quasi-nominal adverb of time (e.g. now, tomorrow)
17. Total adverbs	RA	adverb, after nominal head (e.g. else)
	REX	adverb introducing appositional constructions (e.g., namely)
	RG	degree adverb (very, so, too)
	RGQ	wh- degree adverb (how)
	RGQV	wh-ever degree adverb (however)
	RGR	comparative degree adverb (more, less)
	RGT	superlative degree adverb (most, least)
	RL	locative adverb (e.g. alongside, forward)
	RP	prep. adverb, particle (e.g about, in)
	RPK	prep. adv., catenative (about in be about to)
	RR	general adverb
	RRQ	WH wh- general adverb (where, when, why, how)
	RRQV	WH-EVER wh-ever

		general adverb (wherever, whenever)
	RRR	comparative general adverb (e.g. better, longer)
	RRT	superlative general adverb (e.g. best, longest)
	RT	quasi-nominal adverb of time (e.g. now, tomorrow)
18. total nouns	ND1	singular noun of direction (e.g. north, southeast)
	NN	common noun, neutral for number (e.g. sheep, cod, headquarters)
	NN1	singular common noun (e.g. book, girl)
	NN2	plural common noun (e.g. books, girls)
	NNA	following noun of title (e.g. M.A.)
	NNB	preceding noun of title (e.g. Mr., Prof.)
	NNL1	singular locative noun (e.g. Island, Street)
	NNL2	plural locative noun (e.g. Islands, Streets)
	NNO	numeral noun, neutral for number (e.g. dozen, hundred)
	NNO2	numeral noun, plural (e.g. hundreds, thousands)
	NNT1	temporal noun, singular (e.g. day, week, year)
	NNT2	temporal noun, plural (e.g. days, weeks,

		years)
	NNU	unit of measurement, neutral for number (e.g. in, cc)
	NNU1	singular unit of measurement (e.g. inch, centimetre)
	NNU2	plural unit of measurement (e.g. ins., feet)
	NP	proper noun, neutral for number (e.g. IBM, Andes)
	NP1	singular proper noun (e.g. London, Jane, Frederick)
	NP2	plural proper noun (e.g. Browns, Reagans, Koreas)
	NPD1	singular weekday noun (e.g. Sunday)
	NPD2	plural weekday noun (e.g. Sundays)
	NPM1	singular month noun (e.g. October)
	NPM2	plural month noun (e.g. Octobers)
19. Existential *there*	EX	THERE existential there
20. adjectives	JJ	general adjective
21. infinitives	TO	TO，infinitive marker (to)
22. Direct WH questions	DDQ (not in a subordinate clause)	WH wh-determiner (which, what)
	DDQGE (not in a subordinate clause)	WH wh-determiner, genitive (whose)
	DDQV (not in a subordinate clause)	WH wh-ever determiner, (whichever, whatever)
	RRQ (not in a subordinate clause)	WH wh- general adverb (where, when, why, how)
	RRQV (not in a subordinate	WH-EVER wh-ever

	clause)	general adverb (wherever, whenever)
23. THAT relative clauses controlled by a verb	**CST + V** refer to 'CLAWS7 (C7) POS Tageset	that (as conjunction)
24. THAT relative clauses controlled by an adjective	**CST + JJ** refer to 'CLAWS7 (C7) POS Tageset	that (as conjunction)
25. THAT relative clauses controlled by a noun	**CST + N** refer to 'CLAWS7 (C7) POS Tageset	that (as conjunction)
26. WH relative clauses WH-	PNQS (in a subordinate clause)	WH subjective wh-pronoun (who)
	PNQO (in a subordinate clause)	WH objective wh-pronoun (whom)
	DDQ (in a subordinate clause)	WH wh-determiner (which, what)
	DDQGE (in a subordinate clause)	WH wh-determiner, genitive (whose)
	DDQV (in a subordinate clause)	WH wh-ever determiner, (whichever, whatever)
	RRQ (in a subordinate clause)	WH wh- general adverb (where, when, why, how)
	RRQV (in a subordinate clause)	WH-EVER wh-ever general adverb (wherever, whenever)
27. Possibility modals (can, may, might, could)	VM (by using a cluster tool, the frequency of can, may, might, could is able to be added)	modal auxiliary (can, will, would, etc.)
28. Necessity modals (ought, should, must)	VM (by using a cluster tool, the frequency of ought, should, must is able to be added)	modal auxiliary (can, will, would, etc.)
29. Predictive modals (will, would, shall)	VM (by using a cluster tool, the frequency of will, would, shall is able to be added)	modal auxiliary (can, will, would, etc.)

30. concessive adverbial subordinators (although, though)	CS (although, though)	subordinating conjunction (e.g. if, because, unless, so, for)
31. causative adverbial subordinators (because)	CS (because)	subordinating conjunction (e.g. if, because, unless, so, for)
32. conditional adverbial subordinators (if, unless)	CS (if, unless)	subordinating conjunction (e.g. if, because, unless, so, for)

4.4.1.3 Factor analysis

Once the data is ready, it makes possible starting the statistical computations by using the SPSS program, a Windows based program for a comprehensive and flexible statistical analysis (http://www.spss.co.in/). Since the goal of an MD analysis is to uncover the functionally varied dimensions in a discourse, the factor analysis was used to achieve this. Factor analysis is a type of data reduction technique used to simplify the analysis of a large number of variables by grouping them into a small number of factors on the basis of co-occurrence (technically, shared or common variance) (Kim & Mueller, 1978). In this study, the original variables for the factor analysis are the normalized frequencies of 32 linguistic features. Following Biber's (1988) work, a principal component analysis, or PCA, was adopted for the factor extraction. This method shows the degree of the variance accounted for by each component. Usually, each extracted factor represents the features that co-occur frequently in the corpus, while the first factor extracts the largest group of co-occurring features in the data, maximizing their shared variance. In each factor, two groups of features are usually included: a group of features with high positive factor loadings, and a group of features that have high negative loadings; the presence of each set of feature is associated with the relative absence of the other set (Gorsuch, 1983).

After factor extraction, most factors derived from the extraction phase are usually correlated with many variables; therefore, interpreting any of the factors can be difficult. To solve this problem, I followed Biber (1988) and adopted a rotated solution to transform the initial and

complicated factors into simpler ones, which makes it easier for the interpretation. Specifically, rotating factors forces each linguistic feature to load on as few factors as possible. As a result, “each factor is characterized by the features that are most representative of a particular amount of shared variance” (Biber, 1988, p.82). Rotating factors also enhances the interpretation of the functional dimensions underlying the factors. Therefore, it is more reasonable to assume that the factors extracted from this study are likely to be correlated and become more appropriate for this analysis.

After the factor extraction, a factor or dimension score for each text and for each of the factors was computed. By using SPSS, each text receives a factor score, based on the standardized frequencies of the variables on each factor. The factor score for each dimension was computed by summing up “the frequencies of the features having salient loadings on that dimension” (Biber, 1995, p.117), with the frequencies of the features of the salient positive loadings on a factor or a dimension subtracting that with negative loadings on that factor (by summing both positive and negative). Following Biber (1995), features with salient loadings greater than 0.35 were considered important for the factor score computation. If one feature has salient loadings on more than one factor, only the highest loading was included in the calculation of the score of the factor on which it loaded on, and other loadings on other factors were not considered for factor score calculation. In this way, the distinction among factors and the independence of the factor scores are assured.

4.4.1.4 Determination of the number of factors to retain

Another important issue in the process of factor analysis is determining the number of factors to retain for a group of variables of interest (Fabrigar et al., 1999; Hayton, Allen, & Scarpello, 2004; Velicer, Eaton, & Fava, 2000). This decision is especially important because errors of under-factoring (extracting too few factors) or over-factoring (extracting too many factors) are likely to result in non-interpretable or unreliable factors (Fava & Velicer, 1992; Lee & Comrey, 1979; Wood, Tataryn, & Gorsuch, 1996) and can potentially mislead theory development efforts (Fabrigar et al., 1999). Among three rules for deciding the number of factors to retain (Floyd & Widaman, 1995), i.e., statistical tests, mathematical and psychometric criteria, and

rules of thumb, the eigenvalue-greater-than-one method (Kaiser, 1960) from the *mathematical and psychometric* criteria and the scree test (Cattell, 1966) from the *rules of thumb* was adopted in the current MD analysis. The eigenvalue-greater-than-one rule/Kaiser-Guttman criterion (Kaiser, 1960) is a very widely used method based on population proofs regarding the size of the eigenvalues for uncorrelated variables. In addition, the scree test (Cattell, 1966) is a very practical criterion from the rubric of *rules of thumb,* which explains the percentage of variance accounted for, and the number of variables that have significant loadings on the factor (Floyd & Widaman, 1995; Reise, Waller, & Comrey, 2000).

4.4.1.5 Cluster analysis

In the MD analysis, a cluster analysis was used for the identification of text types that can be interpreted functionally by considering the communicative purposes of the texts grouped into each type. A cluster analysis is a multivariate statistical technique used to classify objects into similar sets or groups (Johnson & Wichern, 1992). By using this approach, text types are identified: the texts grouped together in a cluster are maximally similar linguistically, while the different clusters are maximally different linguistically (Biber, 1995). The clustering of text types is done by using a K-Means Clustering method in SPSS, the first procedure of which is to pre-define the number of clusters.

The determination of the number of clusters is a difficult issue in cluster analysis. Everitt (1979) regarded this fundamental step as among the yet unsolved problems of cluster analysis, and some researchers consider there are no completely satisfactory methods so far for the determination of the number of clusters for any type of cluster analysis (Bock, 1985; Hartigan, 1985). Even though some tests are proposed to justify the number of clusters (Milligan & Cooper, 1985), most of them are poorly understood or are heuristic (Aldenderfer & Blashfield, 1984); they did not have a better performance than *pseudo-F* test (Caliński & Harabasz, 1974), which looks for consensus among the statistics, i.e., local peaks of the pseudo F statistic. The *pseudo-F* variable formulated is presented as:

$$pseudo\text{-}F = \frac{A}{W}\frac{n-k}{k-1}$$

In this formula, where *A* and *W* are the among- and within-cluster sum of squares respectively, *n* is the number of objects, and *k* is the number of clusters. All these can be achieved by using a SPSS software; and this test is considered as guidance for the current study.

4.4.1.6 The exploration of CB text types across industries

Based on the CB text types identified and different industries that the 1020 CB texts are categorized into, the relationship between different CB text types and industries were revealed. The first step I took is to categorize corporate blogs into several industrial groupings properly. It was decided that the industry standards of ISIC[10] and GICS[11] can be adopted for the categorization of corporate blogs, on the basis of the background information of classification schemes, as well as the nature of corporate blogs. Once this step was complete, I looked into the relationship between CB text types and industries by using some basic mathematical calculations. At first, I calculated the total number of CB texts from each text type that are also grouped into industry, on the basis of the counting of the total number of CB texts in each industry, and the total number of CB texts in each text type. Then, I normalized the numbers of CB texts from various industries, by means of counting the percentage of the total number of CB texts from a particular industry to 1020 (the total number of CB texts in the corpus). Finally, I calculated the ratio of CB texts from each text type across each industry to the normalized number of CB texts from a particular industry. It is assumed that a close relationship between a text type and a certain industry may help me interpret the communicative functions and purposes of corporate blogs more profoundly and pertinently.

4.4.2 Data processing in the exploration of lexico-grammatical features in corporate blogs

Following what has been presented in the previous sections, which concerns methodological issues and data processing in Biber's (e.g., 1988, 2006) MD analysis, this section elaborates further how some linguistic features from SFL system,, i.e., the most common verbs and their

[10] International Standard Industrial Classification (ISIC) of All Economic Activities (http://unstats.un.org/unsd/cr/registry/regcst.asp?Cl=27)
[11] the Global Industry Classification Standard (GICS) system (https://www.msci.com/gics)

process types, pronouns, modal verbs, lexical density and grammatical complexity (cf. Section 2.3.2, Section 2.3.3 & Chapter 7) are processed across different CB text types.

4.4.2.1 The most common verbs and their process types

In order to explore the most common verbs and their process types across different CB text types, the following procedures were followed. At first, several directory files were created, each of which includes a set of tagged txt files that corresponds to a number of pure CB texts from the same text type. Once the data were ready, the software AntConc was employed for data processing. By using AntConc, the first step is to open a directory file, and use the Word List feature to generate a list of ordered words. What follows is a sorting process: click on the 'Sort' button, and sort the words generated according to their frequency. After the identification of the most common verb in each text type, what to do next is to judge their process types which are given different functional labels according to their role in a clause. It is assumed that the same process types, whether material, mental, relational or verbal, can be described by different verbs, and the same verb may be assigned by different means of processes.

4.4.2.2 Personal pronouns

The exploration of the use of personal pronouns across different CB text types basically follows the same procedures to those processed in the exploration of the most common verbs across different CB text types. Once the data were ready, and the decision was made as for what type of personal pronouns can be searched for in the current research, which includes *I, me, we, us*, *you*, *he, him, she, her, they, them,* and *it*, AntConc was employed for data processing. By using AntConc, the first step is to open a directory file, and use the Word List feature to generate a list of ordered words. The second step is to search for each kind of personal pronouns, and observe their frequencies displayed. When the frequency counts of these personal pronouns across different CB text types were done, they should be normalized for a more valid measurement. It is assumed that there are possibly some strong contrasts between the use of these personal pronouns, and some pronouns do not occur at all in the corpus.

4.4.2.3 Modal verbs

The exploration of the use of modal verbs across different CB text types follows the same procedures to those processed for the study of personal pronouns. The difference is this analysis focuses on the use of modal verbs from Biber's (1988, p.241) classification, in which a variety of modal verbs are included:

- Possibility modals: can, may, might, could
- Necessity modals: ought, should, must
- Predictive modals: will, would, shall

Although the exploration of the use of each modal verb follows the same procedures presented in the previous section, the investigation of the frequency of three broad modal verb types allows for a further step by counting modal verbs belonging to each broad modal verb type together, and normalizing them to the same standard.

4.4.2.4 Lexical density

Both lexical density and grammatical complexity (cf. Section 2.3.3.3) are concerned with different kinds of intricacy in written and spoken language. Typically, "written language becomes complex by being lexically dense: it packs a large number of lexical items into each clause; whereas spoken language becomes complex by being grammatically intricate" (Halliday, 2004, p.654). As for the measurement of LD in the current analysis, it has followed a refined concept by Halliday (1985) who defines LD as "the proportion of the number of lexical items to the number of running words" (Halliday, 1985, p.64), as can be seen in the following formula:

- LD (%) = The number of lexical words/ Total number of words (tokens) x100%.

The calculation of LD largely relies on a clear distinction between lexical items and grammatical items, which is usually considered as an important task (Ure, 1971; Ure & Ellis, 1977) because "lexis and grammar form one system – lexicogrammar – and the moving between them is a matter of delicacy (Halliday, 2004, p. 43). What has been adopted in the

current analysis is based on the traditional view on lexical items which include three word classes: nouns, verbs and adjectives (Ure, 1971), as well as Halliday's (1985) divide of some adverbs, in which non-grammaticalized adverbs (including all adverbs derived from adjectives) can be classified as lexical items. In order to explore LD across different text types, LD in each text types should be investigated by dividing the number of lexical items (nouns, verbs, adjectives, and all adverbs derived from adjectives) by the total number of words. This can be achieved by using AntConc, which helps count the number of lexical items as well as the total number of words, and following the procedures listed below: 1) a directory file representing a text type was opened; 2）the Word List feature was used to generate a list of ordered words；3）the frequency of nouns was calculated by summing up any POS tagged words initiated with N; 4) the frequency of verbs was calculated by summing up any POS tagged words initiated with V; 5) the frequency of adjectives was calculated by summing up any POS tagged words initiated with J; 6) the frequency of adverbs derived from adjectives was calculated by summing up any adverbs ending in -ly from those POS tagged words initiated with R; 7) the number of lexical words was summed up by adding the frequencies of nouns, verbs, adjectives, and all adverbs derived from adjectives; 8) the total number of words in each text type was calculated; 9) LD was calculated by comparing the number of lexical words to the total number of words. I employed Ure's (1971) conclusion to measure the LD across different CB text types, which claims that a large majority of the spoken texts have a lexical density of under 40%, while a large majority of the written texts have a lexical density of 40% or higher.

4.4.2.5 Grammatical complexity

In terms of the measurement of grammatical complexity across different CB text types, it mainly relies on the construct of T-unit. The T-unit measurement has been developed by Hunt (1965) and considered to be a more reliable measure of grammatical complexity than those previously used, e.g., sentence length, clause length (Crowhurst, 1979). Two specific measures have been especially popular: mean length of T-unit (MLT), which relies on the overall length in words of the T-unit, averaged across all T-units in a text, and clauses per T-unit (C/TU), which relies on the number of dependent clauses per T-unit, again averaged

across all T-units in a text (Biber et al., 2011).

In this analysis, a number of grammatical complexity features will be computed by using a syntactic complexity analyzer[12], which is designed by Lu (2010) and manipulated in python and runs on UNIX-like systems with Java 1.6 and python 2.5 or higher installed.

Table 4.4: The fourteen grammatical complexity measures automated (Lu, 2010, p. 479)

Grammatical complexity features /measures	Code	Definition
Type 1: Length of production unit		
Mean length of clause	MLC	# of words / # of clauses
Mean length of sentence	MLS	# of words / # of sentences
Mean length of T-unit	MLT	# of words / # of T-units
Type 2: *Sentence complexity*		
Sentence complexity ratio	C/S	# of clauses / # of sentences
Type 3: *Subordination*		
T-unit complexity ratio	C/T	# of clauses / # of T-units
Complex T-unit ratio	CT/T	# of complex T-units / # of T-units
Dependent clause ratio	DC/C	# of dependent clauses / # of clauses
Dependent clauses per T-unit	DC/T	# of dependent clauses / # of T-units
Type 4*: *Coordination		
Coordinate phrases per clause	CP/C	# of coordinate phrases / # of clauses
Coordinate phrases per T-unit	CP/T	# of coordinate phrases / # of T-units
Sentence coordination ratio	T/S	# of T-units / # of sentences
Type 5: Particular structures		
Complex nominals per clause	CN/C	# of complex nominals / # of clauses
Complex nominals per T-unit	CN/T	# of complex nominals / # of T-units
Verb phrases per T-unit	VP/T	# of verb phrases / # of T-units

4.5 Summary

This chapter describes the corpus design and methodology of this research project. At first, several issues concerning the design of CBC were discussed, including the aim and the type of the corpus, representativeness, corpus size and sample size, data collection and ethical issues. And then, the methodology for this research project was introduced which includes two main parts: multi-dimensional analysis and an SFL-based approach. Given the limited range of

[12] The analyzer is called the Stanford parser (Klein & Manning, 2003) for parsing the input file and Tregex (Levy & Andrew, 2006) to query the parse trees

variation within the CBC from corporate blog genre, and possibly some unique external environment of companies who have their blogs, a modified multi-dimensional analysis was employed. Following Biber (1988), a factor analysis and a cluster analysis were performed, which is for the identification of the underlying dimensions/co-occurrence patterns and text types respectively. On the basis of the text types identified, a further exploration of the relationship between different CB text types and industries was also performed for establishing a linking between language characterizing certain text types and industries the corporate blogs belong to, and thus providing a more continuous and detailed functional interpretation of CB text types. Next, some linguistic features that can be mapped onto Halliday's (2004) metafunctions, i.e., the most popular verbs and their process types, personal pronouns, modal verbs, lexical density, and grammatical complexity, were compared across different CB text types identified.

Chapter 5: Results of the multi-dimensional analysis

5.1 Introduction

In this chapter, the results of a multi-dimensional analysis of corporate blogs are presented and discussed. This includes several core stages: I first identify the underlying factors (dimensions) by using a factor analysis and present the results in Sections 5.2, and then provide the interpretation of the dimensions in Section 5.3. With the dimensions identified in the factor analysis being used in a cluster analysis, several clusters (text types) are identified, the results of which are presented in Section 5.4 and interpreted in Section 5.5. On this basis, the relationship between CB text types and industries are further explored and identified in Section 5.6. This is followed by an overall discussion in Section 5.7.

5.2 Results of factor analysis

Following the methodological steps described in Chapter 4, 32 linguistic features in the corpus were tagged, counted and normalized. The normalized frequencies of these features were then used in the factor analysis. Based on the initial factor analysis (after promax rotation), a five-factor solution appeared to be optimal. This preliminary determination relies on Kaiser's (1960) greater-than-one criterion and Cattell's (1966) scree-plot criterion in the measurement literature. At first, in order to keep the factors with eigenvalues larger than one, ten factors extracted from the dataset could be retained. As can be seen in Table 5.1, the eigenvalue for a given factor reflects the variance in all the variables, which is accounted for by that factor. The proportion (% of Variance) indicates the relative weight of each factor in the total variance. Usually, the first factor accounts for the most variance in the variables, the second accounts for the next highest amount of variance, and so on. And the cumulative shows the amount of variance explained by a factor plus all of the previous ones. In particular, the first factor explains 8.976% of the total variance, factor 1 and factor 2 account for 17.80% of the total variance, and the first 10 factors account for 55.103% of the shared variance.

Table 5.1: Eigenvalues of the rotated factor analysis

Factors	Eigenvalues	% of Variance	Cumulative %

1	2.869	8.967	8.967
2	2.827	8.833	17.800
3	2.460	7.686	25.486
4	1.547	4.834	30.320
5	1.400	4.376	34.697
6	1.382	4.318	39.015
7	1.364	4.263	43.278
8	1.348	4.213	47.490
9	1.293	4.041	51.531
10	1.143	3.572	55.103

And then, a 'scree plot' approach, which was suggested by Cattell (1966) and followed by Gorsuch (1983) and Biber (e.g., 1995), was adopted for the determination of the best number of factors to be included in the analysis. By using this approach, the number of factors can be determined if there seems to be a natural break in the eigenvalues or if it is clear that the plot becomes flat at a particular point. In the current factor analysis, the factors identified are shown in Figure 5.1, in which there is a sharp break between the fourth and fifth factors and a lesser break between the fifth and sixth factors. And accordingly, a five-factor solution could possibly be optimal.

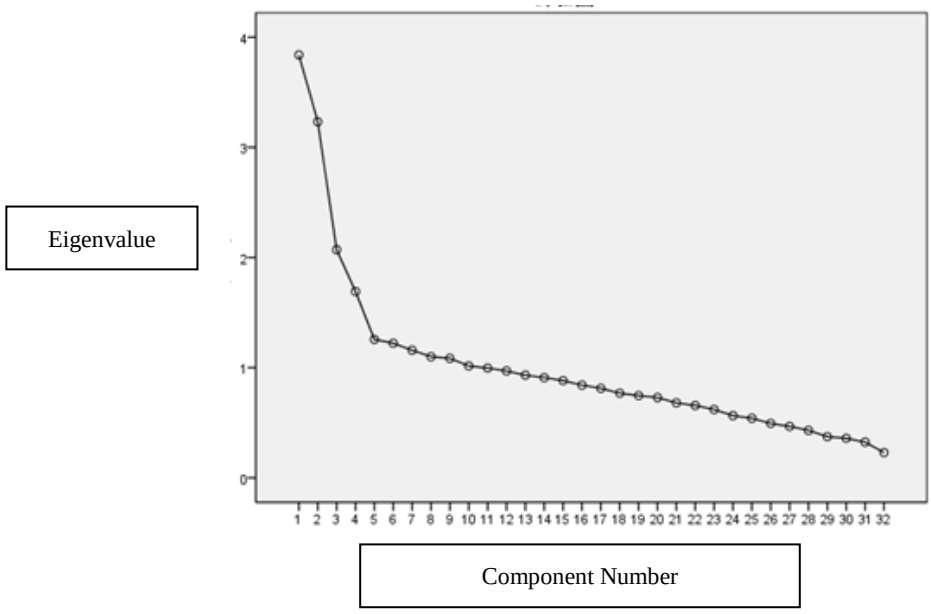


Figure 5.1: The scree plot

However, I consider that this number, which is essentially based on a graphical method, has been underestimated. After a further scrutiny of the original variables in each factor, I decided to add one more factor, i.e., Factor 6, because it comprises a potentially important language

feature: type-token ratio (TTR). TTR is an important criterion to measure the formality/informality and generality/technicality of a text, making it possible to distinguish between general/technical, formal/informal, or spoken/written dimensions. In view of this, if a five-factor model was chosen, an important loading feature would be missed. Although TTR is the only variable/language feature in Factor 6, it is composed of several elements, constituting a multi-layer dimension. Taken together, including six factors for the dataset is more sensible for the subsequent analysis as well as interpretation. Among the factors extracted, the first six factors represent 39% of the shared variance and are readily interpretable. The first factor extracts the largest group of co-occurring features and accounts for the highest shared variance, while subsequent factors account for relatively little additional variance. Once this decision was made, another issue which concerns the inclusion of loading features in the factors was taken into account. In this analysis, I followed the principle set in an early work by Biber (1988), in which only feature loadings having an absolute value of 0.35 or greater are included, as shown in Table 5.2.

Table 5.2: Factor loadings

Factors	Positive/Negative	Linguistic features	Loading
1	positive	First-person pronouns	0.786
		Total adverbs	0.622
		Past tense	0.575
		Third-person personal pronouns	0.534
		Indefinite pronouns	0.457
		WH relative clauses	0.428
		Place adverbials	0.411
		Pronoun *it*	0.397
	negative	Mean word length	-0.606
2	positive	Total nouns	0.812
		Adjectives	0.782
		Total prepositional phrases	0.744
		Infinitives	0.577
	negative	-	-
3	positive	Second-person pronouns	0.764
		Conditional adverbial subordinators (if, unless)	0.685
		Possibility modals (can, may, might, could)	0.618
		Present tense	0.547
	negative	-	-

4	positive	By passives	0.758
		Agentless passive	0.691
	negative	-	-
5	positive	THAT relative clauses controlled by an adjective	0.649
		THAT relative clauses controlled by a verb	0.495
		THAT relative clauses controlled by a noun	0.486
		Perfect aspect verbs	0.476
		Concessive adverbial subordinators (although, though)	0.408
	negative	-	-
6	positive	Type-token ratio	0.778
	negative	-	-

The factor loadings can be used to compute factor scores for the corpus, where factor scores are computed for a text by summing the frequency of the features loading on that factor. On any factors retrieved, the loading of features varies, which indicates the extent to which the feature is representative of a factor. The higher a feature's loading on a factor, the stronger its co-occurrence with the rest of the linguistic features loading on that factor. Features with the highest loadings are thus the strongest predictors of a factor and play an important role in the interpretation of the functions underlying a factor.

What follows is an interpretation of the functions underlying the co-occurring patterns of the linguistic features constituting each factor. It might be difficult to interpret and name the factors merely on the basis of their factor loadings. Rather, the interpretation of a factor could be based on two respects: one involves the analysis of the shared communicative functions of the sets of co-occurring features defining the factor based on a large body of previous studies; the other is the detailed investigation of prototypical corporate blog samples. In short, the analysis of the shared communicative functions provides the basis for factor's functional interpretations and prototypical corporate blog samples help to confirm or refine these interpretations.

5.3 Interpretation of factors

5.3.1 Factor 1: Personal involvement narration

Factor 1 is the most important factor underlying the corpus in this analysis. It accounts for 8.967 % of the total pool of shared variance, and includes the linguistic features with loadings greater than 0.35, as can be seen in Table 5.3. In all, on this factor, the features with positive loadings have a much more stable distribution than the one with a negative loading. Positive features are *First-person pronouns*, *Total adverbs*, *past tense*, *Third-person pronouns*, *Indefinite pronouns*, *WH relative clauses*, *Place adverbials* (*e.g., above, beside, outdoors*), *Pronoun it*. And there is only one negative feature: mean word length. All of these features are important predictors of this factor because none of them has a salient loading on another factor.

Table 5.3: Linguistic features and loadings on Factor 1

Factors	Positive/Negative	Linguistic features	Loading
1	positive	First-person pronouns	0.786
		Total adverbs	0.622
		Past tense	0.575
		Third-person personal pronouns	0.534
		Indefinite pronouns	0.457
		WH relative clauses	0.428
		Place adverbials	0.411
		Pronoun *it*	0.397
	negative	Mean word length	-0.606

Many language features with positive loadings are mainly associated with the first two dimensions in Biber (1988), reflecting such functions as personal involvement and narration. The frequency of first person pronouns, with the highest positive loading, indicates that Factor 1 is associated with a personal style in that first person pronouns could be used for writers to make claims, or expose themselves to the audience (Harwood, 2005), or to claim authority and exhibit some form of ownership of the content (Hyland, 2001, 2002). Other positive features, including pronoun *it,* and indefinite pronouns, are used for substituting for fuller noun phrases and presenting a less precise content. All these features are therefore associated with an interpersonal interaction and the involvement of personal feelings, and could be labelled as

being interactive or involvement-focused, as suggested by Chafe (1982) and Tannen (1985).

It is also of great necessity to take into consideration another set of co-occurring features, which could be interpreted as narration. Past tense and third person pronouns are the two clearest markers of narration. “Past tense verbs are used to describe the major past events that constitute the backbone of the narrative, while third-person pronouns are used to refer to the animate, typically human participants in the narrative” (Biber, 1988, p. 152). Besides, total adverbs and place adverbials that load on this factor are used to make situation-dependent reference for narration. Many of these features, e.g. past tense, adverbs, place adverbials, are a particular type of verbs and adverbs, thus being able to be characterized as a verbal-based style. Such kind of verbal style tends to be associated with narrative texts, which usually relate the unfolding of events and actions in time. Taking into account the two main functions, personal involvement and narration, this factor can be labelled as personal involvement narration. The following text excerpts illustrate the dense use of positive Factor 1 features:

Text Sample 5.1 (Selected positive Factor 1 features are first person pronouns (*italicized*), past tense verbs (underlined), and third person pronouns (**bold**); Factor 1 score = 9.03665; File 18)

> *I* grew up in Woodbridge, VA, a suburb just outside of Washington, D.C. and come from a fabulous, loving and supportive family. My parents were very closely linked to *my* career choice. My mother was a flight attendant for US Air for 30 years, so *I* spent much of my childhood traveling with *my* siblings and learning all about the country and the world first hand. *I* wanted to be a flight attendant just like *my* mother. However, **she** told me *I* needed to be in the cockpit, because that’s where the real fun was. And that, as they say, was that. With the support of my family, *I* found that it’s amazing what you can do when you don’t know that you can’t.
>
> My mom was *my* biggest influence and my biggest fan; **she** was just amazing. **She** had this great job that took her all over the world, and at the same time was always there for all of *us*. **She** was Superwoman. **She** opened my eyes to aviation and, more importantly, the world. Throughout *my* research, **she** actively helped me find a flight school, an instructor and a good college to make *my* dream a reality.
>
> …
>
> The experience *I* have had with Delta Air Lines has been positive from the start; and while *I* have been an employee for more than five years, it has yet to feel like work. When you are here, you really are part of huge family. And despite what the future holds, *I* know that *I* will have the support and help of this new family, Delta Air Lines.

Text Sample 5.2 (Selected positive Factor 1 features are first person pronouns (*italicized*), past

tense verbs (underlined), and third person pronouns (**bold**); Factor 1 score = 8.84835; File 17)

> *We* know you met on a Delta flight—how exactly did it happen?
> *We* met on October 22nd, 2010 while aboard Delta flight 1874 – from Jacksonville to Atlanta. *I* was flying back home to Los Angeles and Chris (my fiancé) was flying to Pittsburgh for a conference. After *we* unsuccessfully tried everything in our power to have *our* seats changed (who really wants to sit in the rear of the plane in row 40!?), *we* found ourselves sitting next to each other. *I* had taken *my* seat first and Chris was one of the last to board. After taking **his** seat, **he** awkwardly removed **his** coat, slightly elbowing *me*. **He** apologized and that was *our* first conversation. As the television monitors were lowered, Chris was quick to make a humorous comment, and *we* laughed a bit. In his 'outdoor' voice, Chris continued to talk over the safety announcements. With headphones in hand, *I* couldn't wait until *we* reached 10,000 ft. so that *I* could politely insert *my* headphones and disengage from conversation with the loud guy sitting next to *me*!
> How did you start talking?
> Before *we* had reached our cruising altitude, *I* found myself engaged in a pleasant conversation with this surprisingly charming neighbor. *We* talked about life, love, friends, religion, and everything in between. After Chris called me out for wearing a faux engagement ring (yes, *I* USED to wear a 'creeper ring' when *I* would fly) and after enjoying a drink on the flight (**he** did not even pay for my drink *gasp*) *we* both discovered that *we* had a 4-hour layover in Atlanta before continuing on to *our* final destinations. After making this discovery, *I* waited for 40A to invite *me* to spend time with **him**. **He** didn't at first; and *I* found *myself* slightly disappointed. Twenty minutes later as the plane started its descent, **he** finally asked if *I*'d like to grab a bite to eat during *our* long layover and of course *I* said 'yes'.

The two text samples presented above are representative of the positive end of Factor 1. Both of them illustrate the combined characteristics of interactiveness, involvement, and narration, characterized by a frequent use of first person pronouns, past tense, third person pronouns, etc.. In text excerpt 5.1, the story ends with a conversational tone by using both the first person pronoun *I* and second person pronoun *you* frequently, which aims to close the social distance between reader and writer. At the beginning of this text sample, past tense, a typical narrative tense, frequently co-occurs with personal pronouns, which is often used to tell us stories through the thoughts of a specific person. Similar to text sample 5.1, text sample 5.2 also exhibits both a narrative and a conversational feature, though in a slightly different way. Text sample 5.2 is a part removed from the CB text named 'File 17', which is written in the form of conversation, consisting of several questions and answers. From the linguistic perspective, it includes several turn-takings, a response marker in dialogue, in which the questions are expectant of some kind of feedback, and have a reply. This text excerpt has turned out to have

a high degree of reply expectation among written texts; the issue of reply expectation has been discussed in Martin (1992, p. 511) who assumed that “communication technology is constantly expanding meaning potential”. Apart from the informal conversational style, this text sample also tells us stories when answering the questions, reflected by a frequent co-occurrence of past tense verbs and personal pronouns. In all, these two samples represent the communicative functions indexed by the sets of co-occurring features on Factor 1.

Besides, there is only one negatively loaded feature on Factor 1: mean word length. Word length marks a high density of information, by reflecting precise word choice and an exact presentation of informational content. Longer words tend to be rarer and more specific in meaning than shorter words (Biber, 1995; Zipf, 1949). With only one feature in the set, the functional interpretation was not attempted.

5.3.2 Factor 2: Informational production

The positive features on Factor 2 are *Total nouns, adjectives, Total prepositional phrases, infinitives*, with no negative features, as presented in Table 5.4.

Table 5.4: Linguistic features and loadings on Factor 2

Factors	Positive/Negative	Linguistic features	Loading
2	positive	Total nouns	0.812
		Adjectives	0.782
		Total prepositional phrases	0.744
		Infinitives	0.577
	negative	-	-

Most of these features are associated with nouns and noun modification: adjectives are noun modifiers, and prepositional phrase often involves a post-nominal structure, thus representing a nominal style. Based on previous studies on differences between spoken and written language by Chafe (1982), Tannen (1982b), Biber (1988), etc., the nominal-verbal dimension is thought of as reflecting the density of information to different extent. A nominal style is said to involve more planning and processing and have a higher density of information. Thus, it is more often found in written instead of spoken texts. In the present study, the functional significance of the grouping of these features tends to be informational-focused (Biber, 1988).

Another linguistic feature *TO* infinitives can be used for a variety of functions, such as integrating or expanding ideas (Chafe, 1985), introducing an aim, goal, and purpose, introducing a method, framing points in a discussion, and introducing a complement (Thompson, 2002). Overall, Factor 2 could be labelled as informational production. The following text samples illustrate a dense use of positive Factor 2 scores.

Text sample 5.3 (Selected positive Factor 2 features are pre- and post-nominal modifiers (*italicized*); Factor 2 score = 10.19942; File 35)

> Read on to find out how you could be one of the *lucky* five.
>
> Start with a *Clean* Slate. Before bringing home your *holiday* groceries de-clutter your pantry and restore order. Here's how:
>
> - Wipe down *dusty* shelves, cans and jars *with warm water* and a bit of dish soap. For *sticky* messes, sprinkle the spot *with baking soda* and top *with a warm, wet paper towel*. Let sit for a minute then scrape off with a spatula. Dry shelves thoroughly before restocking.
> - Place stuff you use every day (cereal, nut butters, snacks) within *easy* reach.
> - It's best to put *lightweight* items on *higher* shelves and *heavy* items on *lower* shelves.
> - Group items together, such as *baking* supplies, *canned* beans and veggies, sauces and soups.
> - Place flours, *dried* beans and nuts in *clear* containers so you can see when you're running low.

Text sample 5.4 (Selected positive Factor 2 feature is *TO* infinitives (**bold**), Factor 2 score = 7.59264; File 9)

> First and foremost, we want you **to know** we're continuing to work with the soldiers individually **to make** this situation right for each of them. We regret that this experience caused these soldiers **to feel** anything but welcome on their return home. We honor their service and are grateful for the sacrifices of our military service members and their families.
>
> …
>
> As an Army wife and 12 year Delta employee I also want **to take** this opportunity **to shed some light on** the additional ways we work **to support** U.S. Military personnel and their families throughout the travel process as well as in the community. We are a longstanding supporter military support organizations such as the Fisher House Foundation and United Service Organization.

Both of the text samples above show a highly informational focused characteristic, reflected by relatively frequent use of *TO* infinitives and pre- and post-nominal modifiers. In particular, text

sample 5.3, which is from "Whole Story" and aims to give some favorite tips concerning how to get pantry in ship shape order and ready for the Thanksgiving feast to their audiences, reflects a dense integration of information. Besides, text sample 5.4 (from "Delta Air Lines Blog") is also information focused, i.e. sharing some information, particularly changes to the official policies, with those who have participated in a recent conversation on baggage allowances for active duty U.S. military personnel. Overall, the text samples shown above represent the communicative functions of this factor.

5.3.3 Factor 3: Interactive persuasion

The positive features on Factor 3 are *Second-person pronouns, conditional adverbial subordinators (if, unless), Possibility modals (can, may, might, could), Present tense,* with no negative features.

Table 5.5: Linguistic features and loadings on Factor 3

Factors	Positive/Negative	Linguistic features	Loading
3	positive	Second-person pronouns	0.764
		conditional adverbial subordinators (if, unless)	0.685
		Possibility modals (can, may, might, could)	0.618
		Present tense	0.547
	negative	-	-

Similar to Factor 1, this factor is characterized by a mixture of co-occurring language feature sets, reflecting several contextual functions. The highest positive loading *second person pronouns* is associated with an interactive style, the characteristics of which are acknowledging the presence of the reader and directly referring to the audience of the text. Smith (1982) has even indicated that the second person pronoun *you* should be regarded as the most interactive form among personal pronouns. Semantically, *you* in the corpus data representing corporate blogs refer to people or target audiences in general (readers), rather than a particular person. A reader-writer relationship is more clearly displayed if the other two features, i.e. possibility modals and conditional adverbial subordinators (if, unless), are considered. Possibility modals can be used to disguise the source of evaluation giving an impression of objectivity (Hyland, 2001; Vold, 2006). The co-occurrence of these two features

is associated with an argumentative/persuasive discourse (Biber, 2006; Santini, 2005). Finally, the high frequency of present tense verbs reflects a focus on current event, in contrast to the past tense verbs that have a positive loading on Factor 1. In all, this factor could be named as 'interactive persuasion'. The following text samples illustrate the set of co-occurring features on positive Factor 3:

Text sample 5.5 (Selected positive Factor3 features are listed below: second-person pronouns are *italicized*, conditional adverbial subordinators (if, unless) are **bold underlined**, and possibility modals (can, may, might, could) are **bold**; Factor 3 score = 5.12889; File 5)

> *You* **can** now purchase Economy Comfort on transcon flights! Check out delta.com for the specific offerings and to book *your* next flight.
>
> Last summer we introduced Economy Comfort on our international flights and it's proven to be a huge hit with all of *you!* Since then we've received a ton of feedback and many of *you* have asked us 'when will *you* make this available on the rest of *your* flights?' Well today we're happy to share that beginning in late spring for travel next summer, *you* **can** select Economy Comfort on more than 800 additional planes!
>
> The product will be offered on domestic and short-haul international flights onboard all of our domestic 767, 757, 737, MD88, MD90, A320, A319, DC9 and all two-class regional jets including the CRJ900, CRJ700, E170 and E175 aircraft. Here's what the seats will offer:
>
> - 34+ inches of pitch. This means even more legroom!
> - Priority Boarding. Now *you* **can** have just a little extra time to stow *your* carry-on and relax as *you* travel to *your* destination.
>
> So how does it work? Once *you* have *your* Economy ticket purchased, *you* **can** choose an Economy Comfort seat via delta.com, our self-service kiosks or Delta reservations for an additional fee of $19-$99 one-way including Alaska, Hawaii, and short-haul international flights. Or, **if** *you* purchased a full-fare Economy class ticket, *you*'ll get complimentary access to the new seats.
>
> Just like our long-haul international Economy Comfort option, **if** *you*'re a SkyMiles Medallion member *you*'ll receive added benefits. Here's the breakdown:

Text sample 5.6 (Selected positive Factor 3 features are listed below: second-person pronouns are *italicized*, conditional adverbial subordinators (if, unless) are **bold underlined**, and possibility modals (can, may, might, could) are **bold**; Factor 3 score = 4.58266; File 458)

> Boy, *you* are all going to be die-cutting experts when I get through with *you*! Today we're going to talk about "coining". **If** *you*'ve been die-cutting for a while, *you* **may** be familiar with this topic already, even **if** you didn't know the technical name. **If** *you* just started using the Fuse Creativity System, then *you*'ll be able to add something new to *your* die-cutting vocabulary!

…

However, **if** *you*'re just using a die only, it's something *you*'ll want to take into consideration. For instance, **if** *you* decide *you*'d like to use the Wavy Square die to make some cool upcycled coasters from old game boards and want a nice rounded edge on the top side, *you*'ll need to place the game board face down on the die. Or, **if** *you* want to cut some shapes out of soda cans, *you*'ll need to place whatever side *you* want to be seen face down before running it through the Fuse. With softer metals, it is possible to take a small brayer and smooth out the edges **if** *you* need to. As always, **if** *you* have any questions, feel free to ask! Just remember "Face down to make it round"!

Both text samples above show an intensive use of interactive and persuasive features. In text sample 5.5, which is from "Delta Air Lines Blog", possibility modal *can* and second person pronoun *you* are the most common co-occurring words. Most of such co-occurrences indicate the function of influencing the attitudes and intentions of the audience, e.g. *You* **can** now purchase Economy Comfort, *you* **can** select Economy Comfort, and *you* **can** choose an Economy Comfort seat. Text sample 5.6 is from "Fiskars Ambassador", a blog by a producer of consumer goods. This text sample shows a frequent co-occurrence of conditional adverbial subordinator *if*, and second person pronoun *you*, providing an overall framing of argumentation. Overall, the text samples shown above represent persuasive as well as interactive functions of this factor.

5.3.4 Factor 4: Abstract /impersonal style

Factor 4 includes two positive features: *By passive*, and *Agentless passives*, as can be seen in Table 5.6.

Table 5.6: Linguistic features and loadings on Factor 4

Factors	Positive/Negative	Linguistic features	Loading
4	positive	By passives	0.758
		Agentless passive	0.691
	negative	-	-

Both of the two features are passive constructions, the most important markers of abstract production which give texts an impersonal style (Biber, 1988; Csomay, 2006; Martin, 1992; Westin, 2002b; Westin & Geisler, 2002). The two passive forms are used to give the entity acted upon into prominence of fronting: the agentless passive is a form that results from the deletion of the agent, and by passive (agentive passives) usually contains agent by-phrases at

the end of the sentence. Overall, texts with frequent use of these features tend to be abstract and impersonal. And the co-occurrences of these features are illustrated in the following text samples:

Text sample 5.7 (Selected positive Factor 4 features are agentless passives (underlined) and by passive constructions (**bold**), Factor 4 score = 6.53686; File 36)

> Homeopathy has a long history of use all over the world:
> •Homeopathy comes from the Greek words homeo meaning "similar" and pathos meaning "suffering." It **was first used by** Dr. Samuel Hahnemann in 1796.
> • Homeopathic medicines **are regulated as drugs by** the FDA and are made according to the specifications of the Homeopathic Pharmacopoeia of the United States (HPUS), which lists approximately 1,280 medicines.
> •If one chops several onions in a small kitchen, one will experience symptoms such as spasmodic sneezing, runny nose and itchy eyes. All these symptoms will **be improved by** breathing fresh air. A patient experiencing these same symptoms, including feeling better when breathing fresh air, whether **caused by** allergy or a cold, **will be relieved by** a homeopathic preparation of the onion also known as Alliumcepa. It is as if a very small amount of the onion helped the body to react better against symptoms of cold or allergy similar to those caused by the onion.

Text sample 5.8 (Selected positive Factor 4 features are agentless passives (underlined) and by passive constructions (**bold**), Factor 4 score = 6.17869; File 534)

> Two days after a Japanese journalist, Mika Yamamoto, was shot and killed in the Syrian city of Aleppo, her news agency released some of the footage she recorded in her final hours.
> The video, *posted online* with subtitles by The Telegraph of Britain, shows that Ms. Yamamoto, 45, was filming Syrian rebel fighters alongside her partner, Kazutaka Sato, when she was shot and killed.
> Video recorded by Mika Yamamoto, a Japanese journalist, in the Syrian city of Aleppo this week, shortly before she was shot and killed.
> According to a biographical note on the Web site of her agency, the Japan Press, Ms. Yamamoto was an experienced war correspondent who produced a report on the oppression of women in Afghanistan during the Taliban's rule, and later reported on the American-led wars in both Afghanistan and Iraq. Colleagues told The Japan Times that she was a careful reporter and "wasn't a reckless type."
> …
> Ms. Yamamoto's death was announced on Monday, when graphic video of her body being transported out of Syria was uploaded to YouTube by supporters of the rebel Free Syrian Army.

The two text samples above are from "Whole Story" and "Lede" respectively, in which a

number of passive constructions are used. Among these passive constructions, some of them are agentless passives, which do not state who carried out the action, and some are by passive constructions. The use of agentless passives can be attributed to several reasons: one is that the agent is not that important or unknown although implied (Csomay, 2006; Quirk & Greenbaum, 1993), as illustrated in the following sample sentences: *Mika Yamamoto was shot and killed* (by someone), *Homeopathic medicines are made* (by someone) *Ms. Yamamoto's death was announced* (by someone); another reason could be the passive expresses a general statement, while the implied agent is 'people' or 'one' (Downing and Locke, 2002), as illustrated in the following sample sentence: *the homeopathic healing concept can be seen* (by you). Besides, by passive construction is also frequently used in text sample 5.7 and 5.8, which intends to put the emphasis on agent rather than the action. In the following sample sentences, the retained agent is typically inanimate, e.g. *will be improved by breathing fresh air, caused by allergy or a cold, will be relieved by a homeopathic preparation of the onion, caused by the onion*, etc.. Overall, the text samples shown above represent the communicative functions of this factor.

5.3.5 Factor 5: Evaluative stance

The positive features on Factor 5 are *THAT relative clauses controlled by an adjective, THAT relative clauses controlled by a verb, THAT relative clauses controlled by a noun, Perfect aspect verbs, concessive adverbial subordinators (although, though)*, with no negative features (see Table 5.7)

Table 5.7: Linguistic features and loadings on Factor 5

Factors	Positive/Negative	Linguistic features	Loading
5	positive	THAT relative clauses controlled by an adjective	0.649
		THAT relative clauses controlled by a verb	0.495
		THAT relative clauses controlled by a noun	0.486
		Perfect aspect verbs	0.476
		Concessive adverbial subordinators (although, though)	0.408
	negative	-	-

The most striking feature on this factor is the use of that-clause controlled by an adjective, a

verb, or a noun, the structure of which provides a means to talk about the information in the dependent clause (Winter, 1982). This structure has an evaluative potential, that is, the authors' stance is given in the main clause, and the propositional information is given in the *that*-clause (e.g., *it is possible that we don't detect...*). Many grammatical devices can be used to express stance; as suggested by Biber (2004b), the *that* complement clauses could be controlled by a verb, an adjective, and a noun respectively. In particular, the evaluative *that* patterns are introduced by a range of different predicates, including non-factive/communication verbs/nouns, attitudinal verbs/nouns/adjective, factive/certainty verb/nouns, and likelihood verb/nouns/adjective[13](Biber, 2004b).What has been adopted in the interpretation of Factor 5 is the classification of evaluative function of reporting verbs by Francis, Hunston, and Manning (1996), which overlaps Biber's (2004b) work to some extent. Referring to the classification of verbs in Francis et al. (1996), several groups of reporting verbs were found in the corpus: SAY group, THINK group, DISCOVER group, ARRANGE group, SHOW and FIND group. The explanations and verb samples that are adapted from Francis et al. (1996, pp. 97-101) are listed below:

A. SAY verbs are concerned with speaking, writing, and other forms of communication, e.g., *boast, claim, complain, explain, say, suggest*.

B. THINK verbs are concerned with thinking, including having a belief; knowing, understanding, hoping, fearing, e.g., *think, assume, feel, hold, believe*.

C. DISCOVER verbs are concerned with coming to know or think something, including verbs which indicate that someone remembers something they had forgotten or not thought about for some time, e.g., *discover, find out, notice, realize, remember*.

D. ARRANGE verbs are concerned with causing something to happen, which includes people making arrangements for the future, or situations determining the future, e.g.,

[13] Sub-categories and samples:
non-factive/communication verb (e.g., imply, report, suggest)
attitudinal verb (e.g., anticipate, expect, prefer)
factive/certainty verb (e.g., demonstrate, realize, show)
likelihood verb (e.g., appear, hypothesize, predict)
attitudinal adjectives (e.g., good, advisable, paradoxical)
likelihood adjectives (e.g., possible, likely, unlikely)
non-factive/communication noun (e.g., comment, proposal, remark)
attitudinal noun (e.g., hope, reason, view)
factive/certainty noun (e.g., assertion, observation, statement)
likelihood noun (e.g., assumption, implication, opinion)

guarantee, mean, arrange, ensure, mean.

E. SHOW and FIND verbs are concerned with indicating a fact or situation or with coming to know or think some-thing, e.g., *show, demonstrate, reveal, find, observe, discover, indicate*.

These verbs usually co-occur with that-clauses which are generally known to index information integration to expand the idea-unit (Biber, 1988). In particular, that-clauses controlled by adjectives, particularly evaluative adjectives, are used to enhance the value of the research reported in the academic paper (Charles, 2006; Hunston & Thompson, 2000). And that-clauses controlled by a noun are used to identify the status of the information presented in the that-clause, including argument, assumption, claim, idea, notion, possibility, and fact (Biber, 2004c). Apart from that-clauses, the use of perfect aspect verbs usually reflects a focus on current events (Li & Ge, 2009). The concessive adverbial subordinators are "subordinating conjunctions, words which introduce (mainly finite) dependent clauses" (Biber et al., 1999, p. 85). In a more general sense, the same subordinators may introduce various types of adverbial clauses, and present clauses with various meaning (Quirk et al., 1985, p. 1077) Overall, the co-occurrence of the positive features is associated with 'stance-focused discourse', framing the authors' evaluation. The following text samples illustrate a dense use of positive Factor 5 features:

Text sample 5.9 (Selected positive Factor 5 features are that clauses controlled by an adjective (*italicized*), that clauses controlled by a verb (**bold**), and that clauses controlled by a noun (**bold underlined**), Factor 5 score = 7.60212; File 778)

> It's *important that* your social media presence is a good one because customers may **find that** before they find your actual website. For example, **a Tweet that** contains your company name may rank higher than your company blog. If that is the case, you want to **make sure that** you are putting you best foot forward.
>
> …
>
> If your social media profile is the **only thing that** you customers can see, you have some problems. **Make sure that** there is a link to your **company website that** is clearly visible. Make a mention of your website to all customers who see your Twitter page. In fact, you can even use it as your quote in your profile. If you use Google+, **make sure that** you have a picture of yourself. This is because your picture will be located next to **every post that** is indexed. Customers are more likely to spend money on your products if they know who you are. When customers do spend money on your products, **make sure tha**t you

offer credit card processing options.

Text sample 5.10 (Selected positive Factor 5 features are that clauses controlled by a verb (**bold**), and that clauses controlled by a noun (**bold underlined**), Factor 5 score = 5.67987; File 313)

> One can find non-legal assistance **declaring that** they can carry out many of the lawful **projects that** are required. Any personal who **believes that** it is too expensive to seek the services of a proper lawful service needs to **realize that** attorneys are part of a consistent occupation, which **means that** there are **rules that** they need to comply with, as well as moral requirements required by the Arizona States Bar. An personal who is not a residence attorney does not also share the same requirements. This is also not the appropriate time for an personal to do elements on his or her own since purchasing a home is often a customer most expensive buy.
>
> While some lawful types utilized in residence may appear the same, buy or folder agreement types change from one state to another.
>
> Additionally, a variety of name insurance companies and companies **require that** one maintains an attorney to **create sure that**, among other elements, the residence is free from liens, it has an obvious name, and that the sale is going to close as expected.

Text sample 5.11 (Selected positive Factor 5 features are that clauses controlled by a verb (**bold**), and that clauses controlled by a noun (**bold underlined**), Factor5 score = 4.28981; File 634)

> At first we were disappointed, because we **know that** all our products are safe by scientific standards and meet or exceed government regulations. Over time, though, we've come to **realize that** sometimes safety alone isn't enough. There's a vigorous public discussion going on around the world about what ingredients should or shouldn't be in personal care products, and how they should be regulated. We have **a point of view that** we've expressed, based on our considered understanding of the science involved, and that's always going to be our starting point. But what matters most isn't what we think, it's what the people who use our products think.

These text samples are from "Marketing Conversation", "Business Poperty", and "JNJ BTW" respectively. All of them illustrate the frequent use of that-complement clauses and relative clauses. These features are typically used to provide information as well as presenting people's attitudes or stance towards the proposition. In text sample 5.9, *It's important* (adjective) *that* is a typical example of extraposition, which allows the writer to push long and complicated chunks of information to the end of the clause. This is able to avoid disturbing the normal pattern of end-weighting in English (Bloor & Bloor, 1995), and to simultaneously

express a stance, an effect which Halliday and Matthiessen (2004, p. 615) call “explicitly objective” modalization, contributing to the presentation of an affective stance in this structure. Most of that-clause instances in three texts are controlled by a verb, such as *find* (the FIND group) *that*, *make sure* (the ARRANGE group) *that*, *declaring* (the SAY group) *that, believes* (the THINK group) *that*, *realize* (the THINK group) *that*, *means* (the SHOW group) *that*, *know* (the THINK group) *that*. They could be divided into different semantic groups according to the grammatical pattern of the verb (Francis et al., 1996). Most of these examples overlap with the attitudinal verb or factive/certainty verb that are supposed to have evaluative potential in Biber (2004b). However, in that-clauses controlled by noun, most of them are primarily used for informational elaboration with an informational focus. The use of *a point of view that* in text sample 5.11 is an exception in that it is an attitudinal stance noun, and used to show a writer’s personal feelings or opinions towards something. Overall, the text samples shown above represent the communicative functions of this factor.

5.3.6 Factor 6: General information vs. technical information

TTR (the type-token ratio) is the only positive feature on Factor 6, as can be seen in Table 5.8. The type-token ratio is a measure of lexical diversity, reflecting the richness of the vocabulary of a text. A high TTR indicates a large amount of lexical variation and a low TTR indicates relatively little lexical variation. In previous studies, it has been identified that TTR is lower in conversation than in the written registers, and somewhat lower in academic writing than in writings of lower technicality (Johansson, 2012; Tagliacozzo, 1976). Overall, this factor could be named as ‘general information vs. technical information’. This interpretation would be more valid with the consideration of prototypical text samples.

Table 5.8: Linguistic features and loadings on Factor 6

Factors	Positive/Negative	Linguistic features	Loading
6	positive	Type-token ratio	0.778
	negative	-	-

Presented below are two text samples from “FedEx Citizenship Blog” and “Infusionsoft Blog” respectively. Text sample 5.12 has a high factor score, while text sample 5.13 has a rather low factor score, which means TTR of two text samples should be different from each other. The

measurement of TTR in two text samples is able to prove this.

Text sample 5.12 (Factor 6 score = 2.18894; File 622)

> As summer temperatures heat up, grills get fired up, kids have fun in the sun and we spend more time with our families and take the vacations we promised them we would. Download this desktop wallpaper below!
>
> A few things to look forward while you kick back and fire up some burgers and hot dogs on the grill:
>
> Infusionsoft's Summer 2012 Release is coming out soon, where we expand the capabilities of the Campaign Builder so users have even more confidence and ease in building their marketing campaigns. Sign up for our webinar on July 9th where we will detail the upcoming enhancements.
>
> We have upcoming live training classes at Infusionsoft University that train you to be successful using the Infusionsoft app. It's perfect for new employees or if you want a refresher yourself.
>
> The Infusionsoft Marketplace is growing! It is the spot to check first if you need an Infusionsoft Certified Consultant to help or you want the perfect integration with Infusionsoft.

Text sample 5.13 (Factor 6 score = -.49882; File 131)

> Most pharmaceuticals need to be stored and transported in one of two temperature ranges: Controlled Room Temperature (15-25°C) or Cold / Refrigerated (2-8°C). Typically the sophistication of the molecules used in the drug and/or stability of it drives the need for tighter temperature control (i.e., more refrigeration).
>
> Other healthcare products such as blood, tissues, cell lines and blood-based products (e.g., customized medicines / biotech drugs) often require a Frozen (-20 down to -80°C) or even Deep Frozen / Cryogenic (-150°C or below) range in order to keep any cellular degradation of the samples/specimens from occurring during transit. When customers ship in those Cold, Frozen or Deep Frozen ranges, however, they use active or passive cooling systems contained within specialized containers or packaging in order to maintain those temperatures (such as dry ice/gel packs in parcel shipments). They do extensive testing in their quality departments to validate and approve these methods, but one area that they often have trouble with is air transport and particularly how temperatures can fluctuate inside of an aircraft. Most of their packaging is validated for a specific external/environmental range and if the temperature dips below that it can cause a Cold shipment to drop below the desired minimum temperature of 2°C and freeze, for example, which can often render the product unusable and it has to be destroyed. The largest segment of pharmaceutical product from a temperature standpoint is in the CRT range, and particularly drugs moving in bulk. For these shipments, especially generics, the manufacturers or distributors are often more price conscious and do not want to incur the added cost of using expensive temp-control packing materials or specialized containers. Instead they opt to ship on days where the temperature is mild at origin and destination and hope that they won't get any temperature excursions in transit. And, even then, the stability of these products is typically good so they can have allowable excursions of a

certain duration down to 5°C and up to 30°C without the product being damaged or unusable.

Based on the computation of their type/token ratio (TTR), it can be viewed that TTR of text sample 5.12 (68%) is higher than that of text sample 5.13 (56%), which means the text samples differ from each other in terms of their formal/informal and academic/non-academic styles. In particular, text sample 5.12 seems to be less academic than text sample 5.13. Text sample 5.12 focuses on conveying information with personal involvement, characterized by a high frequency of personal pronouns, nouns, attributive adjectives, and a more diverse use of vocabulary, while the prototypical text sample 5.13 focuses on the technical topic of the store and transportation of pharmaceuticals, with the use of a restricted technical vocabulary on a narrow subject matter. Overall, these text samples represent the communicative functions indexed by the co-occurring features on Factor 6.

5.3.7 Summary

To sum up, based on a factor analysis of functional linguistic variation across a 590,663-word corpus of corporate blogs, six principal dimensions of linguistic variation were identified, which represent significant patterns of functional linguistic variation for this variety of language. These six principal dimensions are (1) personal involvement narration, (2) informational production, (3) interactive persuasion, (4) abstract/impersonal style, (5) evaluative stance, and (6) general information vs. technical information. They are further used as predictors in a cluster analysis, through which six clusters (text types) are identified, as presented below.

5.4 Results of cluster analysis

The previous section discussed the major dimensions of linguistic variation in corporate blogs, which were extracted using a factor analysis. These dimensions were then used to identify the main text types of the CB variety by using a cluster analysis, a statistical approach for forming groups of similar objects. In this analysis, a k-means clustering method available on SPSS was adopted, and then the pseudo-F test was used to determine the optimal number of clusters. The interpretation of the clusters identified mainly relies on two types of information: the six

dimension scores for each cluster, and detailed consideration of prototypical corporate blogs from each cluster, where corporate blogs closest to a cluster centroid are considered to be prototypical members of that cluster.

In this cluster analysis, six clusters were identified as optimal on the basis of an analysis of peaks in the clustering criterion and the F distribution. While there is no completely satisfactory way to determine the number of clusters, the measure of Pseudo-F statistic is accepted for the present study (Bock, 1985; Everitt, 1979). As can be seen in Figure 5.2, the axis on the left is for the F distribution, and the horizontal is for the number of clusters. Considering that a sudden rise indicates the optimal number of clusters in the data, and Pseudo-F peak at Cluster 6, 6 is accepted as the number of clusters for the present study.

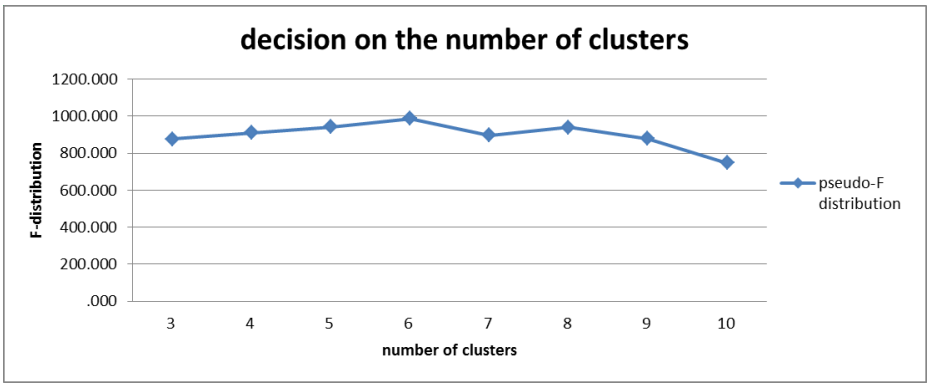


Figure 5.2: Decision on the number of clusters

In the ANOVA table presented below, $F_{calculated}$ is given in the F column, which is computed from the formula F = MSR/MSE: F is the ratio between MSR and MSE, where MSE is called mean square error, and MSR is called mean square regression. Besides, the P-value, or observed significance level of the F-ratio, is obtained in the Sig. column. Since this P-value is smaller than 5%, the result is statistically significant.

Table 5.9: Results of ANOVA

		Cluster		Error		F	Sig.
		Mean Square	df	Mean Square	df		
REGR Factor score	1	52.469	5	.746	1014	70.314	.000
REGR Factor score	2	119.109	5	.418	1014	285.214	.000
REGR Factor score	3	107.534	5	.475	1014	226.538	.000
REGR Factor score	4	105.700	5	.484	1014	218.512	.000

REGR Factor score	5	94.162	5	.541	1014	174.175	.000
REGR Factor score	6	12.147	5	.945	1014	12.853	.000

Besides, Table 5.10 provides the basic descriptive statistics for the six clusters identified by the cluster analysis: the number of corporate blog text in each cluster, nearest cluster, and distance between cluster centroids. These statistics show that the clusters identified differ notably: the smaller clusters tend to be more sharply distinguished linguistically. Cluster 3, which contains the fewest corporate blogs, has a great cluster distance to its nearest cluster. And similarly, cluster 6 has a relatively small number of corporate blogs, but is at a large distance from its nearest cluster. In contrast, the vast majority of the CB texts could be classified into cluster 2, which is the nearest cluster to cluster 4, and also close to cluster 1, cluster 5 and cluster 6. All these seem to show that most text types identified in corporate blogs are close to each other, and thus the linguistic features in different text types may overlap in some way, although there are also some distinctive text types involved in CB, i.e., cluster 3 and cluster 6. And accordingly, the interpretation of text types was assumed to be a more challenging task.

Table 5.10: Cluster analysis results

Cluster	Frequency (total=1020)	Nearest cluster	Distance between Cluster centroids					
			1	2	3	4	5	6
1	120	2		2.308	9.403	2.818	2.769	5.756
2	525	4	2.308		9.386	1.988	2.001	5.391
3	4	6	9.403	9.386		9.260	9.527	6.822
4	160	2	2.818	1.988	9.260		2.550	5.615
5	193	2	2.769	2.001	9.527	2.550		5.575
6	18	2	5.756	5.391	6.822	5.615	5.575	

This descriptive summary is followed by a description of mean dimension scores across clusters, which could be taken into account for choosing values to represent relatively poor and relatively better off clusters. As can be seen in Table 5.11, the average factor scores vary notably across six clusters. Cluster 3 is characterized by an extremely large positive score on Factor 1 and 2, reflecting a narrative and informational focused discourse. Similarly, cluster 6 also has very large positive scores on Factor 2, reflecting a frequent use of features associated with informational production. Cluster 1, Cluster 4 and Cluster 5 are characterized by their relatively lower factor scores in general. In spite of this, the highest use of features

related to positive scores on Factor 4 reflect an abstract style in Cluster 1, the highest use of features associated with positive scores on Factor 5 reflect evaluative stance in Cluster 4, and the highest use of features related to positive scores on Factor 3 reflect an interactive persuasion in Cluster 5. At the other extreme, cluster 2 is even more unmarked in its factor scores, in that each factor score is small and negative.

Table 5.11: The mean dimension scores across clusters

	Cluster analysis					
	Cluster1 (descriptive discourse)	Cluster2 (technical conversation)	Cluster3 (interactive story telling)	Cluster4 (evaluative information report)	Cluster5 (persuasion)	Cluster6 (interactive instruction)
Factor score 1 (personal involvement narration)	-.30313	-.04264	7.20388	-.23089	.24105	1.13125
Factor score 2 (informational production)	.00403	-.02529	5.02429	-.06188	-.45477	5.02053
Factor score 3 (interactive persuasion)	-.33104	-.40018	-.34650	-.29498	1.45641	.96207
Factor score 4 (abstract/impersonal style)	1.92302	-.36000	1.42104	-.22236	-.05190	-.10307
Factor score 5 (evaluative stance)	-.24608	-.35061	2.21280	1.52317	-.20590	.04333
Factor score 6 (general information vs. technical information)	-.03127	-.19989	-.77939	.41275	.22011	.18282

5.5 Interpretation of clusters

5.5.1 Cluster 1: Descriptive discourse

Cluster 1 is relatively common (120 CB texts grouped into this cluster or 11.76% of all CB texts in the corpus), and has the highest positive score on Factor 4 (abstract/impersonal style), as well as a small positive score on Factor 2 (informational production). CB texts in this cluster exhibit an impersonal-style linguistic feature while conveying the information. Presented below

is a prototypical text sample from this cluster.

Text sample 5.14. (Positive Factor 4 features are marked by **bold**; Positive factor 2 features are marked by *italics*; Distance to centroid: 1.113; File 137)

> FedEx was the:
> *First transportation logistics company* to push for *commercial vehicle fuel economy/greenhouse gas legislation and regulation*.
> *First U.S. transportation logistics company to set* a *global aviation reduction carbon dioxide goal*.
> *First U.S. transportation logistics company to set* a *commercial vehicle fuel efficiency fleet goal*.
> *First transportation logistics company to launch* the *first purpose-built electric delivery vehicles*.
> *First transportation logistics company to certify* its corporate headquarters to Leadership in Energy and *Environmental Design*, and made Certification the standard for *newly-built US facilities*.
> And, according to *an environmental non-governmental organization*, FedEx "led the launch that changed the marketplace" with regard to hybrid electric commercial trucks.
> …
> This *reforestation project* is establishing *commercial forests* at Uchindile and Mapanda districts in the *Tanzanian Southern Highlands* in an area that **was classified** as degraded grassland.
> Thailand
> *The Rachathewa landfill* is Thailand's *first sanitary landfill*. Waste **is isolated** from the environment until it has completely biodegraded biologically, chemically, and physically — meaning harmful leachate discharge **is prevented**. The *landfill gas collection system* at Rachathewa generates renewable energy by extracting the methane that occurs during this *remediation process* and uses it to generate *clean electricity*. The electricity **is sold** to the *Metropolitan Electricity Authority*.
> Turkey
> The Mare Manastir Wind Farm **was installed** in the Turkish province of Izmir to provide *renewable power* to *the Turkish electricity grid*, thereby displacing some of the need for non-renewable natural resources **to be used** for the generation of power.

In this prototypical text sample, some salient linguistic features from Factor 2 (informational production) and Factor 4 (abstract/impersonal style) are dominantly used. Those marked by **bold** are passive constructions from Factor 4, which is often characterized as being formal and impersonal. The formality of passive patterns is consistent with the distribution among registers, with high frequencies in academic prose and with conversation at the opposite extreme (Biber et al., 1999). It is difficult to judge what registers might be included in this text

type, since many different registers have been proved to be marked by a high frequency of passive patterns, such as academic discourse, and news. In spite of this, what can be sure is the purpose of the use of passive construction is to leave the doer of an action, and place given information in subject position, giving rise to a focus on information rather those unknown or unimportant people/things. Apart from the language features from Factor 4, those from Factor 2, such as nominal groups including nouns and noun modifications (*italics*), are also dominantly used in this text sample. The use of such language features is assumed to reflect the density of information to different extent (Biber, 1988; Tannen, 1982a). To some extent, these informational-focused language features relates to the distinction between what is given or new in a discourse (Chafe, 1976; Halliday, 1967). As for CB texts from this cluster, including the prototypical text sample presented above, they tend to provide some new and factual information, rather than some already shared knowledge. Both literary and expository/descriptive texts may serve to add new information, and this cluster could be better named as a 'descriptive discourse'.

5.5.2 Cluster 2: Technical conversation

Cluster 2 is much more common (525 CB texts grouped into this cluster or 51.47% of all the texts in the Corpus). This cluster has the largest negative score on Factor 3 (lack linguistic features associated with interactive persuasion), Factor 4 (lack linguistic features associated with abstract/impersonal style), Factor 5 (lack linguistic features associated with stance-focused); a negative score near 0.0 on Factor 1 (personal involvement narration) and Factor 2 (informational production). At the same time, this cluster has a large negative score on Factor 6 (general information vs. technical information).

This cluster, consisting of the largest number of CB texts, has a relatively unmarked language feature in general: neither being too persuasive, abstract, nor being stance-focused. Different from those language features on factors from 1 to 5, TTR from Factor 6 is a feature occurring on a bi-dimensional plane, showing dynamic connectivity between spoken & written registers, and academic & less academic writing (Johansson, 2012; Tagliacozzo, 1976). A low negative score on Factor 6 does not mean this cluster lacks any marked features concerning TTR.

Rather, TTR involves many language features associated with spoken & written registers, and academic & less academic writing, e.g., pronouns, terminologies. Both a technical and conversational/informal style could be noted in this text type, based on a detailed exploration of many CB texts from this text type. Some text samples are presented below:

Text sample 5.15 (Selected features are pronouns (**bold**), terminologies (*italicized*); Distance to centroid: 0.499; File 86)

> Good afternoon everyone. Thank **you** for joining us. Two years ago, Toyota and Tesla Motors announced the beginning of a collaborative effort to bring an all-electric *RAV4* to market this year. Tesla brought to the table a fast and flexible development style…. Toyota shared its expertise in engineering and manufacturing. The idea was to learn from each other. **We** accelerated our conventional development process without compromising product quality. The Toyota/Tesla team continually focused on the customer experience…. How do **we** deliver an unconventional product to mainstream customers that is compelling, affordable, and convenient? This unique collaboration allowed the talents, and creativity, of both Toyota and Tesla to shine. Every aspect of the vehicle, from the handling to the headlamps, was designed, engineered and executed to ensure Toyota quality. A typical product planning cycle is more than four years. The *RAV4 EV* moved from conception to production in less than two. There was no template for this project…..no guidelines….just a challenge to bring a premium EV to the market. The Toyota and Tesla engineers both rose to that challenge. The fruit of this unique partnership is a truly ground-breaking vehicle…. Ladies and gentleman, may I present…….the all new *RAV4 EV*…….
>
> It's all about blending the best of two worlds. The all new *RAV4 EV* marries the efficiency of an *EV* with the versatility of a small *SUV*. Tesla's high performance battery and power train meets the style, *cargo capacity* and performance of the *RAV4 V6*. In fact, it will be the only all-electric *SUV* on the market. And it brings with it an expected driving range of approximately 100 miles – thanks to a uniquely blended *regen-brake system.* The 41.8 kWh *lithium-ion battery pack* can smoothly and quietly take **you** from 0-60 in 7 seconds. That is seriously fast. Of course, the downside to all this electric power is additional weight. The RAV4 EV weighs approximately 470 pounds more than its *V6* counterpart. But the engineers turned this challenge into another opportunity for improvement. By shifting the extra weight downward and towards the center of the vehicle, they achieved a lower center of gravity……comparable to a conventional sedan. The chassis was tuned to support this new weight distribution, creating an extremely comfortable ride. The engineers' original goal was to match the ride and handling of the *RAV4 V6*. The result was an even better driving experience…that I believe…will dramatically change **your** thinking regarding EV performance. On the outside, engineers focused on clean, aerodynamic performance. The RAV4 EV features a distinctive upper and lower *grille, front bumper, side mirrors and spoiler.* The new *LED headlights, tail lights and daytime running lights* combine energy efficiency with sleek lines.
>
> …

> There are many people responsible for creating this fantastic product, but I would like to recognize two individuals, whose leadership was instrumental…. First, our project leader… the *RAV4 EV* chief engineer …..Mr. Greg Bernas.. And **we**'d also like to welcome and thank Tesla Motor's Chief Technology Officer…… J.B. Straubel. Gentlemen, thank **you** for joining us today. J.B., **we** understand Elon wanted to be here today, but **he**'s a little busy. In fact, **he**'s at mission control right now launching a rocket that will dock with the international space station.....because that's what **you** do on a Monday afternoon. Please give **him** our best.

Text sample 5.16 (Selected features are pronouns (**bold**), terminologies (*italicized*); Distance to centroid: 0.569; File 1052)

> **I**'m spending a couple of humid days in Miami Beach. While the beach actually does look nice, the opportunity in Latin America looks even better. Aside from **my** own comments, **I**'ve been listening to bankers from the region, a speaker from Microsoft, Business News Americas, BNY Mellon.
>
> …
>
> The views are consistent and interesting – the economies are fairly robust in the region, averaging around 4% *GDP* if I recall. The returns from the banks are reasonable, with *ROEs* over 10%, a figure that most US and European banks would die for. The *regulatory pressures* are as real as elsewhere. A banker from Ecuador told me that there are NO fees allowed in the *credit card* business (that is 'none'), so the only way to make money in that business is via spreads and *transaction fees* to the SEs. Not much different in Peru and some of the other markets. Interestingly over 1/2 of the population remains unbanked, while the takeup of mobile technologies is high growth. This must clearly lead to taking share via mobile phone and tablet, not by the usual routes to market. Some good examples of use of Social Media as well – Banamex' Facebook page caught my attention. **I** look forward to learning more tomorrow, but this continues to be an attractive market. No surprises there for the Spanish – heck, **I** think Santander and BBVA earn over half their profits from Latam.

Both of text sample 5.15 and 5.16 are prototypical texts with a close distance to the centroid. As illustrated in text sample 5.15, an author from Toyota shares the expertise in engineering and manufacturing of auto industry with his/her readers in an interactive writing/speaking style. Correspondingly, the language features display a combination of both technical and conversational/informal style: frequent use of terminologies, acronyms, personal pronouns, and topic shifts from greeting to information report (Wardhaugh, 1985). In text sample 5.16, the author seems to comment on the views on economic issues from other bankers. As expected, some terminologies and acronyms are used, and simultaneously, a very interactive style is applied at the beginning of the text by means of telling personal events for the purpose of

informing and entertaining readers (e.g. *I'm spending a couple of humid days in Miami Beach. While the beach actually does look nice…*).

Overall, this cluster consists of the largest number of CB texts, and shows a very interactive style in the discussion of technical issues. All these CB texts tend to be an initiator in a conversation, inviting readers to share their comments and opinions, rather than a one-way conversation. Taken together, the name of this cluster could be 'technical conversation'.

5.5.3 Cluster 3: Interactive story telling

Cluster 3 is the most specialized, with the fewest number of texts (only 4, or about 0.39% of the CB texts in the corpus). Linguistically, these texts are extremely personal involved and narrative (Factor 1) and informational focused (Factor 2). This interpretation of Cluster 3 would be more valid if samples from blogs that are prototypical of Cluster 3 are considered, as listed below:

Text sample 5.17 (Positive factor 1 features first person pronouns (**bold)** and third person pronouns (**bold underlined)** are selected; Distance to centroid: 3.074; File 18)

> **My** name is Dawn Cook and **I** am a First Officer for Delta Air Lines. Currently in **my** fifth year with Delta, **I** fly the Boeing 737 (800/700). **I** am based in the New York area, so **I** fly out of JFK, LaGuardia and Newark a lot. In addition to **my** usual routes to the west coast (LAX, SEA, etc.), **I** have also been flying to South America a lot recently. What **I** love about flying and working for Delta is that every flight and every day is different, meaning new opportunities and challenges to do **my** best. Regardless of the number of times **I** have taken off or landed at a certain airport, this job requires that **I** am constantly working for perfection, never allowing **my** professionalism to take a back seat.
>
> **I** grew up in Woodbridge, VA, a suburb just outside of Washington, D.C. and come from a fabulous, loving and supportive family. **My** parents were very closely linked to **my** career choice. **My** mother was a flight attendant for USAir for 30 years, so **I** spent much of **my** childhood traveling with **my** siblings and learning all about the country and the world first hand. **I** wanted to be a flight attendant just like **my** mother. However, **she** told **me I** needed to be in the cockpit, because that's where the real fun was. And that, as **they** say, was that. With the support of **my** family, **I** found that it's amazing what you can do when you don't know that you can't.
>
> **My** mom was **my** biggest influence and **my** biggest fan; **she** was just amazing. **She** had this great job that took **her** all over the world, and at the same time was always there for all of **us**. **She** was Superwoman. **She** opened **my** eyes to aviation and, more importantly,

the world. Throughout **my** research, **she** actively helped **me** find a flight school, an instructor and a good college to make **my** dream a reality. As a result of **her** encouragement and support, **I** was even able to start college with a Private Pilot certificate, a year ahead of most of **my** classmates.

Text sample 5.18 (Positive factor 1 features first person pronouns (**bold)** and third person pronouns (**bold underlined)** are selected; Distance to centroid: 3.473; File 33)

To inspire your efforts, here are a couple of Thanksgiving tales from our very own Team Members:

L'aura:

"Nearly two decades ago, a bunch of Team Members (most of **us** without local family) decided to celebrate Thanksgiving at someone's house in Palo Alto. Because of the holiday, the drive down from North Bay on a Thursday was unnaturally quiet and peaceful. **I** entered a house full of people and food, including eight different kinds of pie and the first turducken **I** had ever seen. After dinner **we** went for a walk around the neighborhood to enjoy the beautiful autumn weather. Back at the house, a co-worker got out **his** guitar and **we** lounged around the living room listening to music and talking till the sun went down. **I**'m thankful for the memory of such a peaceful day spent with like-minded souls in a beautiful part of our country — and that **I** work with such friendly people who welcome all like family."

Liz:

Years ago two friends and **I** had just moved to Galveston, Texas, where **we** rented a beach house. Thanksgiving was approaching, so **we** each invited one new co-worker for dinner along with **their** partners. The plan was seven of **us** for dinner. Just before our guests were due to arrive **I** went to the phone booth across the street to call home. A dog followed me and on the way back **I** met a man without dinner plans, so **I** brought him and the dog with **me**. Eight for dinner. Moments later a car filled with six people stopped to ask directions. **I** didn't know where they were headed but asked if **they** wanted dinner, and if so to come on in. **They** did. Fourteen. One roommate brought two more people **he** met on his walk on the beach, and the dog invited a cat. **We** had a lovely celebration with a patchwork of new faces.

Who and what do you remember from the Thanksgiving you are most thankful for?

Tell **us** in a few short sentences in the comments section below by November 24 for a chance to win a $50 Whole Foods Market gift card.

If you're still in the sharing mood, stop by our Holidays hub and share your holiday snapshots with the Whole Food Market community and show **us** how you celebrate. While you're there make sure to check out our festive recipes, cooking tips and entertaining guides.

Both of the text samples illustrate a dense use of 'narrative' and 'informational' features, thus display a very strategic and intimate communication between writer and reader through

storytelling. In text sample 5.17, the author introduces him at the beginning, and then narrates a series of events in his personal and working life. In text sample 5.18, the reader was informed of a contest to be held, and then a series of thanksgiving tales were told from his/her very own Team Members. In the end of this text, the author seems to initiate a conversation by means of asking a question and a frequent use of second person pronoun *you.* This text type therefore does not correspond exactly to the traditional text types (Werlich, 1976); rather, it could be named as 'interactive storytelling'.

5.5.4 Cluster 4: Evaluative information report

Cluster 4 is also relatively common (160 texts, taking up 15.69% of the CB texts in the corpus). This cluster has the highest positive Factor 5 score and a moderately high positive Factor 6 score, reflecting a stance-focused characteristic, and a non-academic style. At the same time, Cluster 4 is characterized by negative features on Factor 1 (lack linguistic features associated with informational production), Factor 3 (lack linguistic features associated with interactive persuasion), and Factor 4 (lack linguistic features associated with abstract/impersonal style), and a score near 0.0 on Factor 2.

The following text samples illustrate a dense use of positive factor 5 features, including that clause controlled by adjectives, verbs and nouns (**bold**), and a high positive factor features, i.e. a very rare use of technical words.

Text sample 5.19 (Positive factor 5 features are marked by **bold**; Distance to centroid: 0.953; File 245)

> In the 70 page **report that** was issued by the FCC's EAAC, Recommendation T2.2 **stated that** "The EAAC **recommends that** the FCC remove the requirement for TTY (analog real-time text) support for new IP-based **consumer devices that** implement IP-based text **communications that** include, at a minimum, real time text or, in an LTE environment, IMS Multimedia **Telephony that** includes real-time text." But before we can remove that requirement, we need to make texting to emergency services a reality.
> Several trials have been launched in cities around the United States. Black Hawk, Iowa, Durham, North Carolina, and most recently the state of Vermont have all implemented **pilot programs that** allow users to send text messages to the 911 centers.
> Although this is encouraging, when you read the fine print, it's hard to determine what

they're actually testing, and there seems to be quite a few **restrictions that** users should be aware of. For example, in the Vermont trial, users are **cautioned that**:

…

This is why the **trials that** exist here in the US are limited. They use 911 as a pointer to a five digit short **code that** terminates at a specific PSAP. Location isn't sent, and there are questions as to if call back is even possible. They are typically restricted to a single PSAP, and a single carrier. So although I applaud the industry for trying to move the stake forward, at the same time I fear a false sense of security could be instilled in the general public who doesn't have the technical background to digest what's actually happening. I think we can all agree, sending a text message to a five digit short code works just fine. Haven't we **proved that** over the last decade of American Idol results?

This is where the App may be able to play an important role. We've talked before about FRESS, Share with 911, and Smart 911. These are all great ideas, but what they lack is a common infrastructure at the PSAP. In a way, they are all proprietary in their client to server communication. That makes me **believe that** maybe it's time for some redirection of effort. If the App truly solves the problem, and all three of those companies **believe that** it will, let's standardize the back-end so we don't accidentally create a proprietary pairing of software between the public user and public safety.

Text sample 5.20 (Positive factor 5 features are marked by **bold**; Distance to centroid: 1.028; File 889)

The headline actually misrepresents the research findings as they are reported in the body of the article: the research does not **show that** right-wingers are less intelligent, it **shows that** less intelligent people are right-wing (which does not exclude the **possibility that** intelligent people might also become right-wing). Or more precisely, it is rather that less intelligent children tend to gravitate towards right wing politics as adults. Still, it seems reasonable to **assume that** some of these simple-minded children have managed to get a decent education and upbringing, to the **point that** they can now formulate complex right-wing strategies and sometimes even defend them with a convincing level of articulacy.

Of course, accurate descriptions rarely make dramatic headlines, and it seems **likely that** the Mail has deliberately chosen more inflammatory phrasing in order to produce the very controversy it claims to describe. Among my own social media connections, I see Mail Online articles being visited and recirculated (and read?) as much by left-wing readers as right-wing ones. The Mail seems happy here to prioritise being entertaining and shocking over being ideologically coherent, and in this sense, this article is perhaps smarter than some gave it credit for, if it is increasing traffic and active participation from left- and right-wing readers alike. This is arguably a **strategy that** makes more sense online than in the hard copy, where the core who are willing to pay for its content are likely to be more ideologically homogenous.

The use of the labels “left” and “right” is also cunningly misleading. I would **argue that** it is not possible to give a definition of the left/right divide without already thereby positioning oneself somewhere on the left or right, and the definition deployed by the article (and

research?) is both bizarrely left-wing and infuriatingly bland. “Left-wing” here appears to be a label slapped on **opinions that** I would identify more as “liberal” – tolerant of other cultures and the socially marginalised. What is missing is the most meaningful distinction between left and right, the economic one. For me, a right-winger is someone who prefers a small role for government, with market forces giving signals for the allocation of resources and citizens managing “their own” wealth. A left-winger prefers their government to take a more active role in the provision of public services and the redistribution of wealth.

In the two text samples presented above, most of the that-clauses are controlled by nouns (e.g. *report, devices, communications, programs, restrictions, possibility, point, strategy, opinions*), a small number of them are controlled by verbs (e.g. *stated, recommends, cautioned, proved, show, assume, argue*), and one controlled by adjective (*likely*). On the whole, the nouns occurring before that-clauses label the status of the information presented in that-clauses (Biber, 2006). In particular, among these nouns used in the two text samples above, some of them demonstrate a particular stance on a particular topic which depends on some previously established statements/facts, e.g. *report, possibility, point, strategy and opinions*; however, some other nouns, such as *devices, telephony, programs,* etc**.** are not closely related with the realization of viewpoints of the author. In terms of the use of that clause controlled by adjective, *it seems likely that* is also a typical example of extraposition, which simultaneously expresses a stance, an effect which (Halliday & Matthiessen, 2004, p. 615) called “explicitly objective” modalization, contributing to the presentation of an affective stance in this structure.

Among the verbs following that-clauses, most of them are concerned with inter-personal communications. For example, based on the categorization of verb meaning groups in Francis et al. (1996), many of them can be categorized as the SAY group, including the verb **stated**, which indicates the function of saying something in a way that shows a person’s attitude, the verb **recommends**, which is concerned with putting forward a suggestion; the three verbs, the verb **caution**, which means warning someone about problems or danger, the verb **prove**, refers to providing evidence by giving examples, explanations, etc., the verb **argue**, which concerns presenting reasons and argument. Besides, **assume** belongs to the THINK group, and **show** is obviously a SHOW group verb.

The frequent use of these verbs or nouns introducing a that-clause shows the prevalence of a stance construct; however, a further look at the context in the sentences shows that, in most cases, they intend to present information objectively, rather than expressing a personal view point subjectively. The following examples show this, which involves activity or fact engaged in by inanimate agents: 'the research does not *show that'*, **'**Recommendation T2.2 *stated that'*, etc.. Overall, texts from this cluster are primarily informational, in spite of a frequent use of stance adjectives, verbs and nouns. They mostly present information objectively which involves activity or fact engaged in by animate or inanimate agents, instead of showing a personal viewpoint subjectively. Taken together, this cluster could be named as 'evaluative information report'.

5.5.5 Cluster 5: Persuasion

Cluster 5 is relatively more common (193 texts, taking up 18.92% of the CB texts in the corpus). This cluster has the highest positive score on Factor 3 (interactive persuasion), as well as a small positive score on Factor 1 (Personal involvement narration) and Factor 6 (general information). At the same time, cluster 5 is characterized by the largest negative score on Factor 2 (lack linguistic features associated with informational production), a moderate negative score on Factor 5 (lack linguistic features associated with stance-focused), and a score near 0.0 on Factor 4 (abstract/impersonal style).

As illustrated in the following text samples, some Factor 3 linguistic features, such as possibility modal (**bold**), conditional adverbial subordinators (**bold underlined**) are marked.

Text sample 5.21 (Distance to centroid: 0.464; File 228)

> Starting today, **if** you live within 45° North or South of the Equator, we **can** deliver a fresh EC2 server to you in 15 minutes or less. This is a genuine, physical server. We've launched (literally) some brand new technology in order to make this a reality. Read on to learn a lot more.
>
> There are two delivery modes: terrestrial and atmospheric.
>
> Terrestrial Delivery
>
> **If** you live in a densely populated urban area, a uniformed delivery person will have your new server on your doorstep in a matter of minutes. As I write this, trucks loaded with servers are circling the 100 largest cities in the country. Here's one of our delivery people in action:

Atmospheric Delivery

The Atmospheric Delivery model is a lot more interesting. In conjunction with our friends at NASA JPL, we've launched a fleet of satellites in to low Earth orbit. Each satellite is stocked with a considerable number of Cluster Compute Eight Extra Large (cc2.8xlarge) servers, individually packaged in our proprietary re-entry shields.

When you order a server (currently limited to one per customer) using the new Deliver Instance button, we'll select a satellite and place your order in the appropriate delivery queue. After a set of careful (checked, double-checked, and then re-checked) ballistic calculations, the satellite will release your order on a trajectory that will deliver it to the latitude and longitude of your choice, accurate to a 1 meter radius, within 10 minutes. You need do nothing more than fill out this dialog: So far so good, right? Read on, it gets even better!

…Because the server includes a built-in Wi-Fi card and a preconfigured Elastic IP Address, you don't have to connect any cables. You **can** simply leave it where it lands and start using it. In fact, under optimal conditions, you **can** start using it while it is still decelerating. You'll be up and running in minutes.

When you are done with your server, you **can** initiate the return process via a single click in the AWS Management Console.

Text sample 5.22 (Distance to centroid: 0.480; File 868)

As most readers know, my dog Darcie was over 18 years of age when she passed away a few years ago. For an 18 year old she was pretty healthy. She **could**n't see too well. She **could**n't hear much and she didn't like to take long walks. But hey, **if** I'm doing as well as Darcie was when I'm 126 I'll be very happy.

Darcie's biggest problem was probably chronic pain from arthritis in her cervical spine. She **could**n't go down steps very well so she didn't sleep in our bedroom with us anymore. She **could** still go up steps, just not down because going down involved extending her neck to see the steps and that was too painful for her.

To have a veterinarian diagnose and treat her problem would probably have involved several hundred dollars of radio-graphs, MRI's, other lab tests and the expense of lifelong pain medication.

Too bad Darcie didn't have insurance.

…

You **can** get a Purina Care policy as early as 8 weeks of age. **If** you get a policy that covers preventive care everything you spend on vaccinations, lab tests, exams, de-worming, spaying or neutering, flea and heart worm prevention, applies to the deductible. In many cases spending on these routine procedures will completely satisfy the deductible and **if** something weird happens requiring veterinary intervention, you'd be covered for that amount less your co-pay.

Puppies and kittens by definition should have few **if** any pre-existing conditions. And with Purina Care, **if** your puppy or kitten has an hereditary condition that has not yet shown clinical signs, you should be covered for that as well.

The example I like to use for hereditary conditions is hip dysplasia. Many large breed dogs carry the genes for hip dysplasia but they **may** not start showing clinical signs of

pain or lameness until they are over 2 years of age. Some **may** never show signs, but unfortunately many do and in some cases they require surgery and long term care.

If you had insured your large breed dog as a puppy before he showed any signs of hip dysplasia and kept your policy in force, you'd be covered for this condition when it was subsequently diagnosed.

Insurance is about peace of mind and being in a position to provide the best possible veterinary care when and **if** you should need it without having to worry so much about cost. That peace of mind is worth a lot to some people.

Both of the text samples above show a general tone of persuading audiences to buy something: text sample 1 aims to persuade their audience to buy a new product EC2 server, and text sample 2 seems to persuade their potential customers to buy insurance for pets. However, the two samples look like helping the audience, rather than helping the company to make a sale, and accordingly, this cluster may not be classified into a traditional persuasive/argumentative text type.

The two prototypical text samples distinguishes from traditional persuasive/argumentative text type, in that it is not merely by means of reasoning and using evidence, but by other measures, such as establishing credibility (e.g. *Read on to learn a lot more.*), emotion (e.g. *So far so good, right? Read on, it gets even better!*) in text sample 5.21 and sharing personal stories (*e.g. most readers know, my dog Darcie was over 18 years of age when she passed away a few years ago…Too bad Darcie didn't have insurance.*) in text sample 5.22. In general, CB texts in this cluster tend to achieve a traditional persuasive communicative purpose by using a more complex and varied communicative approach. Thus, this cluster could be named as 'persuasion'.

5.5.6 Cluster 6: Interactive instruction

Cluster 6 is relatively specialized (with only 18 texts grouped into this cluster, taking up 1.76% of the CB texts in the corpus) in comparison with other text types. Linguistically, this text type is especially marked for being highly informational (Factor 2), relatively interactive & narrative (Factor 1), and somewhat persuasive (Factor 3). As illustrated in the following text samples, some typical linguistic characteristics, particularly some nominal features, including nouns and noun modification (*italicized*), and personal pronouns (**bold underlined**) are marked.

Text sample 5.23 (Distance to centroid: 1.271; File 21)

Almond butter is *a thick, rich, delicious alternative* that stands perfectly well on its own. **You** can buy it roasted, salted, crunchy, smooth, salt-free, raw and sprouted. What a variety! It can be used in any recipe that calls for *peanut butter* and is wonderful across the board from *sweet favorites* to *savory sauces*. It's a *culinary fact* that *almond butter* makes a *healthy, tasty addition* to *breakfast, snacks, meals, sauces, baked goods, dips and desserts*.

If **you**'ve never tried almond butter or have only had it on bread, **you**'re in for a real treat!

- Stir it into *hot cooked oatmeal* or other *favorite hot breakfast cereal*.
- Spread it over toast and eat plain or top with *fresh or dried fruit*.
- Make a dip with *almond butter* and *cream cheese*; use it with *fresh cut fruit* and *veggie sticks*.
- Make *almond butter cookies*. **You** will love these *No-Bake Thumbprint Cookies*. Here's our very own *Whole Wheat Almond Butter Cookie recipe* and here is our recipe for *Cocoa-Oat Truffles*.
- Make *almond butter* and *banana sandwiches*, or make these *Apple Granola Sandwiches*.
- Add to smoothies like this *Cherry Almond favorite* and this *Chocolate Almond Dream Smoothie*.
- Use in **your** favorite *peanut butter recipes*.
- Use in place of or along with *sesame tahini* when making hummus.
- Spread over pancakes, French toast and waffles. Here's a recipe for *Almond French Toast*.
- Make frosting with *almond butter* and a *little honey and cinnamon*; try it on *muffins and cupcakes*.
- Try a sandwich of *almond butter*, *sprouts* and *thinly sliced red onion* (one of my personal favorites)!
- *Almond butter* makes this *Winter Squash Crostini* so good!

...

For something new and different, try almond butter in place of peanut butter in these recipes:

- *Lettuce Wraps* with Noodles
- *Chocolate Monkey Shake*
- *Vietnamese Spring Rolls*
- *Coconut Curry Sauce*
- *Chocolate Chip Cookies* with Fleur De Sel

For **you** *peanut lovers* with no *peanut problems*, **I** say enjoy **your** (natural, non-hydrogenated) *peanut butter*, but why not add some *almond butter* to the mix? **You** can even combine the two for a delicious, hearty spread.

Have **you** tried *almond butter*? Got a *favorite snack* or recipe? **I**'d love to know!

Text sample 5.24 (Distance to centroid: 1.913; File 40)

My current "best ever" recipe has gotten a little simpler:

4-5 *medium avocadoes* (ripe)

1/3 cup *cherry tomatoes* (grape or regular variety, interior pulp removed and sliced into small pieces)

1/4 *small red onion* (finely chopped)

1 *hot pepper* (jalapeño, seeds removed and finely chopped)

2 cloves of *garlic* (minced)

1/4 *teaspoon ground cumin* (or to taste)

1/4 *teaspoon chili powder* (or to taste)

Juice of 1/2 *grilled lime* (sliced in half and grilled in a skillet or on the barbeque)

Salt and *pepper* to taste

Combine all the *ingredients* except the *avocadoes* and set aside for the *flavours* to merge. Set aside the *pits* and roughly mash the *avocadoes* with a *fork*, taking care to leave whole *chunks*. Fold in the rest of the *ingredients* carefully and correct the *seasoning* to taste. Return the *pits* to the *bowl* and serve. Enjoy!

As always **I** enthusiastically encourage input — if **you** have a *family favourite recipe* or *special process* that works, please share. **My** *guacamole* is better than ever but the "'best" is yet to come.

All the text samples above are characterized by a mixture of a variety of communicative purposes, i.e., mainly informational focused, as well as interacting and persuading. In text sample 5.23, cooking with almond butter is advised by introducing its benefits and extensive uses, and in text sample 5.24, the author not only introduces the production of avocadoes in the world, but also gives some advice on how to select avocadoes in the market, and recipes of it. Both of the text samples are interactive, characterized by a frequent use of first- and second-person pronouns and questions, and persuasive that is meant to influence the way the reader thinks or acts.

Overall, texts from this cluster are primarily informational, clearly presenting some facts in a number of lists with an interactive, narrative and somewhat persuasive style. Considering that all these information provided mainly concerns actually telling someone what to do and how to do something, this cluster could be named as 'interactive instruction'.

5.5.7 Summary

With six dimensions produced by the factor analysis being used as predictors in a cluster analysis, six clusters/text types were identified: (1) descriptive discourse, (2) technical conversation, (3) interactive storytelling, (4) evaluative information report, (5) persuasion and

(6) interactive instruction. The result of the cluster analysis shows that the corporate blog is a hybrid genre, representing a combination of various text types. These identified text types, which are relatively fuzzy in comparison with traditional text types, coordinate with each other and work together to achieve different communicative purposes, such as providing high-tech information to their readers/customers in a conversational tone, establishing community through storytelling, helping solve problems by giving instructions and explanations, and new product development or even direct sales to wide audiences through persuasion, etc. Taken together, all these communicative purposes share a common attribute: using corporate blogs as a communication tool for achieving their organizational goals, i.e., the overall goal of advertising, marketing and public relations. In the following section, a further exploration of CB text types across different industries is provided, taking into account more social circumstances in the text-context system, thus adding to our understanding of the influence of contextual features on corporate blogs.

5.6 CB text types and industries

5.6.1 Classification of industries

Measuring how business firms are classified is often critical in research since every empirical study requires good data to draw valid statistical inferences, including the present study. Therefore, it is of great necessity to categorize corporate blogs into industrial groupings properly, before exploring the relations between CB text types and industries. In order to achieve this, International Standard Industrial Classification (ISIC) of All Economic Activities (cf. Appendix 3), and the Global Industry Classification Standard (GICS) system (cf. Appendix 4) have been adopted. This is based on the background information of classification schemes, as well as the nature of corporate blogs. The combination of the two systems for the industrial classification of corporate blogs in the present study should be applicable in that the two classification systems cover a broad range of industries, and even complement with each other, although empirically measuring relatedness of them is difficult, because they do not provide direct links to industry information. Both of GICS and ISIC consist of the classification codes/divisions assigned to the companies. In particular, GICS is a four-tiered, hierarchical

industry classification system: 10 Sectors, 24 Industry Groups, 68 Industries, and 154 Sub-Industries. GICS classifications can be presented in text or numeric format. The full GICS classification for each company is an 8-digit code with a text description. The hierarchical design of the 8-digit coding system allows for easy transition between GICS tiers (Barra, 2008). Presented below is an example:

- Sector: Consumer Discretionary (GICS code: 25)
- Industry Group: Consumer Services (GICS code: 2530)
- Industry: Hotels, Restaurants & Leisure (GICS code: 253010)
- Sub-industry: Casinos & Gaming (GICS code: 25301010)

Besides, in ISIC, the category at the highest level is the section level, which is complemented by the use of a purely numerical system at the division (2-digit), group (3-digit) and class (4-digit) levels (Rev, 2008). Presented below is an example:

- Section C Manufacturing
- Division 10 Manufacture of food products
- Group 107 Manufacture of other food products
- Class 1071 Manufacture of bakery products
- Class 1072 Manufacture of sugar

Taking into account the dynamic nature of corporate blogs which come from different industries, as well as some useful descriptions of industries and sub-industries, the following rules were followed during the course of classification:

1) In many cases, a company is engaged in a wide range of activities, and thus possibly the main line of work in a company can overlap and fall into many different industries. In the current study, it is decided that only the primary codes, which classify the company's most relevant industry and function, rather than the secondary codes, which classify some less important industries, or the ones overlapping from its main line of business, are adopted. For example, Hu Yoshida is a blog for HITACHI, a Japanese company that provides a wide range of business products in the areas of urban infrastructure, information technology, electronic materials, etc., and some business consulting services. This company provides not only the products but the services, and could possibly fall into Section J Information and communication, Section C Manufacturing Division 26 Manufacture of computer, electronic and

optical products, and Division 73 Advertising and market research, etc. It is ultimately decided that HITACHI's most relevant industry is manufacturing, and thus it is categorized into the industry of manufacturing.

2) If corporate blogs launched by the company could be categorized into different divisions/groups at a lower hierarchy level of an industry category but the same section/sector at the highest level, they would be assigned the code/name at the highest level and categorized into the same industry. For example, the blogs, including Brand Channeler, ComScore Voice, Corporate eye, IAGblog, etc. could be categorized into Section J Information and communication Division 63 Information service activities, and The Unboxed Thoughts Blog belongs to Section J Information and communication Division 61 Telecommunication; all of them could be assigned a more global code name 'Information and communication'.

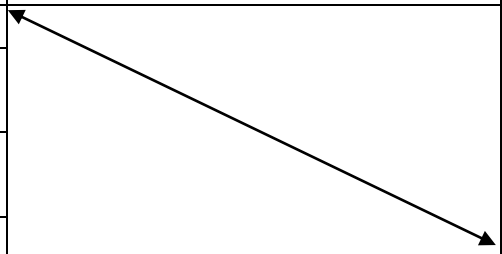
3) A complementary rule is accepted and followed, which means if a blog cannot be categorized into any sections/sectors and divisions/groups in GICS, then it is assigned a name from ISIC, and vice versa. For example, some blogs are coded from the divisions of GICS, such as Business Property (40. Financials 4040 Real Estate), and some blogs are coded from ISIC, such as Delta Air Lines Blog (H. Transportation and storage).

4) If a corporate blog from a company does not fit into any specific classifications in GICS and ISIC, then assign it a name in accordance with its common characteristics. For example, in the data sources used in the current study, the blogs from Facebook Blog, Amazon Web Services Blog, Fandango Movie Blog, the LinkedIn Blog cannot be assigned an industry code in GICS and ISIC, because all of them are concerned with trading in products or services using computer networks. Therefore, I name them 'e-commerce and online social networking service'.

What follows is the industrial classification of corporate blogs (Table 5.12). In this table, both a general and a detailed classification of corporate blogs are provided. In the column of the classifications of Industries, those coded with both two-digit and four-digit numbers are from the divisions of GICS, and some other classifications coded with a capital letter, or a capital letter and a two-digit number are from ISIC.

Table 5.12: The industrial classification of corporate blogs

Names of corporate blogs		Classifications of Industries	More general classifications
1)	Delta Air Lines Blog	H. Transportation and storage	Service industry
3)	Marriott on the move	I. Accommodation and food service activities/25 Consumer discretionary 2530 Hotels restaurants & leisure	
7)	FedEx Citizenship Blog	H. Transportation and storage	
17)	Business Property	40. Financials 4040 Real Estate	
29)	JNJ BTW	Q. Human health and social work activities	
35)	PurinaCare Pet Insurance Blog	40 Financials 4020 Diversified Financials	
2)	Whole Story	G. Wholesale and retail trade/ 30 Consumer staples 3010 Food & staples retailing	Wholesale and retail trade
9)	Turkey Hill	G. Wholesale and retail trade/ 30 Consumer staples 3020 Food, Beverage & Tobacco	
37)	Tyson Foods Hunger Relief	G. Wholesale and retail trade/ 30 Consumer staples 3020 Food, Beverage & Tobacco	
4)	The Cleanest Line	C. Manufacturing	Manufacturing
5)	Toyota open road blog		
18)	Clorox - Dr. Laundry		
22)	Dr. Laskys Blog		
23)	Fiskars Ambassador		
25)	Garmin Blog		
26)	Hu Yoshida		
33)	Rubbermaid - Adventures in Organization		
6)	Red Cross Chat	U. Activities of extraterritorial organizations and bodies	Activities of extraterritorial organizations and bodies
10)	Lede	J. Information and communication 58 publishing activities	Publishing industry
11)	Wallet	J. Information and communication 58 publishing activities	
12)	Comment is free |	J. Information and communication 58	

	guardian.co.uk	publishing activities	
8)	Facebook Blog		E-commerce and online social networking service
13)	Amazon Web Services Blog		
24)	Fandango Movie Blog		
36)	The LinkedIn Blog		
15)	Brand Channeler	J. Information and communication 63 Information service activities	Information and communication
19)	ComScore Voice	J. Information and communication 63 Information service activities	
20)	Converge-A Bader Rutter Blog	J. Information and communication 63 Information service activities	
21)	Corporate eye	J. Information and communication 63 Information service activities	
27)	IAGblog	J. Information and communication 63 Information service activities	
28)	Infusionsoft Blog	J. Information and communication 63 Information service activities	
31)	Marketing Conversation	J. Information and communication 63 Information service activities	
34)	RS Consulting blog	J. Information and communication 63 Information service activities	
38)	The Unboxed Thoughts Blog	J. Information and communication 61 Telecommunications	
40)	Verizon-Policy Blog	J. Information and communication 61 Telecommunications	
14)	Avaya - The Blog	45 Information Technology 4520 Technology Hardware & Equipment	Information Technology
16)	BreakingPoint Labs Blog	45 Information Technology 4510 Software & Services	
30)	Logos Bible Software Blog	45 Information Technology 4510 Software & Services	

32)	Oracle Blogs	45 Information Technology 4510 Software & Services	
39)	Upside Learning Blog	45 Information Technology 4510 Software & Services	
41)	Xerox Blogs	45 Information Technology 4520 Technology Hardware & Equipment	

It is shown in Table 5.12 that the 41 corporate blogs are able to be categorized into 8 industries: service industry, wholesale and retail trade, manufacturing, activities of extraterritorial organizations and bodies, publishing activities, e-commerce and online social networking service, Information and communication, information technology. Most of the corporate blogs can be classified into any divisions or sections in GICS or ISIC, apart from the four blogs named 'Facebook Blog', 'Amazon Web services Blog', 'Fandango Movie Blog', and 'the Linkedin Blog'. All of these four blogs are representative of a new trend of e-marketing: 'Amazon Web services Blog', and 'Fandango Movie Blog' are from a very new and fastest growing industry called e-commerce, which means trading over the Internet; 'Facebook Blog' and 'the Linkedin Blog' are online social networking sites, functioning like an online community of internet users.

5.6.2 CB text types across industries

The exploration of the relationship between CB text types and industries can be achieved by answering the following questions: 1) How do CB text types cut across industries? 2) What are the ratios of the number of CB texts from each text type that are grouped into each industry to a normalized number of CB texts from a particular industry?

In order to answer the first question, the total number of CB texts from each text type that are also grouped into industry was counted, based on the counting of the total number of CB texts in each industry, and the total number of CB texts in each text type, as can be seen in Table 5.13. A full list of CB texts categorized into clusters/text types as well as industry groups is presented in Appendix 5.

Table 5.13: The amount of corporate blogs in 6 text types and 8 industries

Industries	Total	Clusters					
		1	2	3	4	5	6
Service industry	149	17	81	2	18	24	7
Wholesale and retail trade	74	11	34	2	7	9	11
Manufacturing	232	24	121	0	25	62	0
Activities of extraterritorial organizations and bodies	28	4	18	0	3	3	0
Publishing industry	32	19	4	0	7	2	0
E-commerce and online social networking service	101	14	43	0	15	29	0
Information and communication	249	23	124	0	61	41	0
Information technology	155	8	100	0	24	23	0
Total	1020	120	525	4	160	193	18

Table 5.13 shows that the CB texts categorized into six text types can be further grouped into industry groups as well. Overall, it is not an even distribution. The previous cluster analysis shows that, among the 1020 CB texts, there are 120 from text type 1 (descriptive discourse), 525 from text type 2 (technical conversation), 4 from text type 3 (interactive story telling), 160 from text type 4 (evaluative information report), 193 from text type 5 (persuasion), and 18 from text type 6 (interactive instruction). Besides, there is also an uneven distribution of CB texts across industries: there are 149 CB texts from service industry, 74 from wholesale and retail trade industry, 232 from manufacturing industry, 28 from activities of extraterritorial organizations and bodies, 32 from publishing activities, 101 from e-commerce and online social networking service, 249 from information and communication, and 155 from information technology. On these bases, the amount of CB texts from each text type that are grouped into each industry was counted. Taken together, most CB texts from text type 1 (descriptive discourse) are from manufacturing industry, and the industry of information and communication. This is the same case to text type 2 (technical conversation), text type 4 (evaluative information report), and text type 5 (persuasion). In order to further explore the relationship between CB text types and industries, the following computations were performed.

At first, different numbers of CB texts from various industries were normalized, by means of counting the percentage of the total number of CB texts from a particular industry to 1020 (the total number of CB texts in the corpus), as can be seen in Table 5.14. Among the total 1020 CB texts, those from the industry sector of Information and communication makes up the largest percentage, followed by those from the manufacturing industry, and then the industry of information technology, etc.; the blogs from the industry sector of the activities of extraterritorial organizations and bodies make up the small percentage of the total number of CB texts.

Table 5.14: The percentage of the total number of CB texts from a particular industry to 1020 (the total number of CB texts)

Industries	Total	Percentage (normalized number of CB texts from a particular industry)
Service industry	149	14.61%
Wholesale and retail trade	74	7.25%
Manufacturing	232	22.75%
Activities of extraterritorial organizations and bodies	28	2.75%
Publishing activities	32	3.14%
E-commerce and online social networking service	101	9.90%
Information and communication	249	24.41%
Information Technology	155	15.20%

With the percentage of the total number of CB texts from a particular industry to 1020 as a normalized number, the ratio of the number of CB texts from each text type across each industry (Table 5.13) to the normalized numbers (Table 5.14) was counted, as can be seen in Table 5.15.

Table 5.15: The ratio of CB texts from each text type across each industry to the normalized number of CB texts from a particular industry

	Clusters					
	1	2	3	4	5	6
Service industry	116.38	554.50	13.69	123.22	164.30	47.92
Wholesale and retail trade	151.62	468.65	27.57	96.49	124.05	151.62
Manufacturing	105.52	531.98	0.00	109.91	272.59	0.00
Activities of extraterritorial organizations and bodies	145.71	655.71	0.00	109.29	109.29	0.00
Publishing industry	605.63	127.50	0.00	223.13	63.75	0.00
E-commerce and online social	141.39	434.26	0.00	151.49	292.87	0.00

networking service						
Information and communication	94.22	507.95	0.00	249.88	167.95	0.00
Information Technology	52.65	658.06	0.00	157.94	151.35	0.00

In this table, it is shown that in each cluster/text type, the ratio of CB texts from that cluster/text to the normalized total number CB texts from a particular industry is varied. There is always one cluster whose value associated with a particular industry is much higher than others, including 605.63 from both cluster/text type 1 and publishing industry, 658.06 from both cluster/text type 2 and the industry of information technology, 27.57 from both cluster/text type 3 and the industry of wholesale and retail trade, 249.88 from both cluster/text type 4 and the industry of information and communication, 292.87 from both cluster/text type 5 and the industry of e-commerce and online social networking service, 151.62 from both cluster/text type 6 and the industry of wholesale and retail trade. These computations indicate a surprisingly close relationship between certain text types and industries. CB texts from text type 1 (descriptive discourse) predominate in the publishing industry. CB texts from text type 2 (technical conversation) relatively proportionally distribute in different industries, holding a narrow majority in the industry of information technology, and a minimum number in the publishing industry. Among the CB texts from text type 3 (interactive storytelling), most of them are from the wholesale and retail trade industry rather than from the service industry. CB texts from text type 4 (evaluative information report) have a relatively great number of distribution in the industry of Information and communication & publishing industry. For CB texts from text type 5 (persuasion), they predominate in the industry of e-commerce and online social networking service. At last, the vast majority of CB texts from text type 6 (interactive instruction) are from the whole sale and retail trade industry.

The close relationship between certain text types and industry groups shows that there is a close tie between linguistically similar texts which represent a communicative function (Biber, 1989) and specific business functions categorized as operations which equate with the industry code/group (Brown, 2008). That is, in a certain industry group which includes a number of companies, there is a dominant communicative function. Usually, business communications take a variety of forms and serve a number of functions. When businesses use communication strategies well, they are more successful and reap greater rewards from

their efforts. Therefore, the dominant communicative functions associated with specific business functions could be considered as a relatively more plausible and successful strategy. In terms of the business function of a company (e.g., administration, marketing, finance, human resources, information technology, etc.), it is a complex issue: a business function usually consists of multiple activities and process steps (e.g., the process from production, to sales & aftersales service) and a single business process can reflect multiple business functions (Brown, 2008). Without considering this complexity and putting it simply, the current research only focuses on the identification and the interpretation of the relationship between communicative function and the business function. This close linking also reflects the unique as well as complex communication challenges companies from different industries have to face and try to overcome. For example, the close relationship between text type 3 (interactive storytelling) and text type 6 (interactive instruction) and the whole sale and retail trade industry indicates the nature of the business of sales and the strategies adopted. That is, if it is difficult to provide global messages by means of encompassing all of the products in the course of selling, the sellers can either use social media to tell 'stories' that indirectly promote their businesses, without getting into the details of individual parts or products, or providing some instructions about certain products/services. In contrast, a relatively dominant use of persuasive language in the industry of e-commerce and online social networking service means that the blogs from this industry group have used the social media to influence the potential customers' purchasing decisions very directly. Overall, all these varied communicative functions associated with the wider contextual business functions constitute a complex network whose components are varied in the corporate blog.

Based on the results presented above, it is of great necessity to offer a discussion involving both the MD analysis, which encompasses factor analysis and cluster analysis, and the exploration of the relationships between text types and industries, as presented in the next section.

5.7 Discussion

In this section, the results of the multi-dimensional analysis are further discussed taking into

consideration salient characteristics of great variations that emerged from this analysis, which form the identity of corporate blogs. Guided by the research questions proposed in Chapter 1, this discussion starts by summarizing some key findings, and then goes beyond that by showing how this work resolves existing questions in the literature and explaining how it connects to the literature.

By using Biber's (1988) multi-dimensional approach, six factors/dimensions and six clusters/text types were identified. Six factors/dimensions include (1) personal involvement narration, (2) informational production, (3) interactive persuasion, (4) abstract/impersonal style, (5) evaluative stance, and (6) general information vs. technical information. Six clusters/text types include (1) descriptive discourse, (2) technical conversation, (3) interactive storytelling, (4) evaluative information report, (5) persuasion and (6) interactive instruction. On this basis, the relationship between text types and industries was explored, and it was found that there is a close relationship between them. They are text type 1 (descriptive discourse) and the publishing industry, text type 2 (technical conversation) and the industry of information technology, text type 3 (interactive storytelling) and the wholesale and retail trade industry, text type 5 (persuasion) and the industry of e-commerce and online social networking service, text type 6 (interactive instruction) and the whole sale and retail trade industry.

The results of these analyses indicate the complexity and interrelatedness within the language of corporate blogs in terms of the groupings of some sorts of dimensions, text types, and industries. To begin with, six dimensions consistently emerging from MD analyses became candidates for universal parameters of register variation, reflecting the linguistic and functional properties of corporate blogs. On all dimensions, the corporate blogs cover a broad spectrum, with a series of co-occurring linguistic features reflecting a wide range of communicative functions. For blog readers, they may deal with texts with a great information density, as well as some interactive and involved spoken texts. They may encounter texts relying on situated reference, and texts with features of overt persuasion or in an impersonal style. They may also be exposed to some personal stories, as well as some high-tech information. This complexity does not mean that corporate blogs have a quite chaotic communicative environment; rather,

their communicative functions are on a continuous scale of variation with spoken and written registers at two opposite ends. This characteristic was, to a large extent, in good agreement with previous studies. A comparison between the dimensions identified in the current research and the previous MD studies (Biber, 1988, 1995; Biber & Kurjian, 2006) revealed the remarkable similarity between them. Some dimensions identified in the current MD research were largely identical to those in the Biber's 1988 MD study of register variation in English, which provides the fullest account of multi-dimensional methodology and a synchronic analysis of the relations among spoken and written registers. More specifically, 5 dimensions from the current MD analysis, i.e., personal involvement narration, informational production, interactive persuasion, abstract/impersonal style, evaluative stance, resembles those in Biber (1988), in which three dimensions (Dimension 1: Involved versus informational production; Dimension 3: Elaborated versus situation-dependent reference; Dimension 5: Abstract versus non-abstract style) were considered to be indicating sharp distinctions between spoken and written registers, while the other two dimensions (Dimension 2: Narrative Discourse; and Dimension 4: Persuasion) were thought of as having no systematic relationship to spoken and written language (cf. Section 2.3.4.2).

Following the groundbreaking work in 1988, some later MD studies were conducted and similar dimensions were also identified. For example, Dimension 1 (Personal involvement narration) in this study is very close to the first dimension (Personal, involved narration) identified in Biber and Kurjian (2006); the informational production in Dimension 2 is also the positive or negative feature in Biber (1995), Biber and Kurjian (2006), and Grieve et al. (2011); Dimension 4 (abstract/impersonal style) is close to Dimension 5 (abstract versus non-abstract style) in Biber (1995), and Dimension 4 (abstract/technical discourse) in Biber and Kurjian (2006). These resemblances suggest that these dimensions are close to each other in terms of their co-occurring features involved, their commutative functions, and the types of registers they belong to.

A closer look at these resemblances may give rise to an argument on whether or not this comparative work is reasonable since the use of this approach by other researchers (e.g.,

Biber, 1995; Grieve et al., 2011) as opposed to my use of it is based on different data, sampling, and even interpretation. For example, the linguistic features that are representative of informational production on Dimension 1 (Involved versus informational production) in Biber (1988) are *nouns*, *prepositions*, and *adjectives*, *word length* and *type-token ratio*, while the positive features on Dimension 2 (informational production) in the current study include *total nouns, adjectives, total prepositional phrases, infinitives.* It is sure that the current study is different from Biber's MD approach in terms of data, sampling and interpretation. Biber's 1988 work deals with a large number of separate language features, bases his results on statistical procedures, and interpret the results on the basis of a large number of earlier studies. These background studies include previous comparisons of spoken and written texts (Chafe & Tannen, 1987), functional studies of particular linguistic features (Altenberg, 1984; Thompson, 1983), and descriptive grammars of English (Quirk et al., 1985). Therefore, the fundamental principles held in various MD studies can be robust enough to cope with minor differences between data and sampling extracted in different studies. Accordingly, it is unnecessary to focus on minor differences between counts in different data and sampling, as long as effort has been made to present the results in such a way that they are also rooted in a large number of previous studies, and can be reproduced and tested by others. Apart from Biber's work in 1988, his studies in 1995, 2006 and 2011 were also fully referred to. This is due to the fact that the analysis of English registers in Biber (1995) has been either based on the results of Biber (1988), using the LOB Corpus and London-Lund Corpus, or incorporated the findings of several earlier studies, e.g., Biber and Finegan (1989a, 1992, 1994a), Biber, Finegan, and Atkinson (1994), Biber, Finegan, Atkinson, et al. (1994), and Atkinson (1992, 1993). Besides, it is worthwhile comparing web registers (Biber & Kurjian, 2006), blogs (Grieve et al., 2011) and corporate blogs, since all of them is part of the online presence, and thus some attributes should be commonly shared. Overall, these previous studies were reliable, and comprehensive enough for the discussion of dimensions and text types in corporate blogs.

Apart from being categorized into spoken or written registers, the corporate blog is also characterized by a hybrid property, with several feature sets and underlying functions co-occurring in a single dimension. Dimension 1 (personal involvement narration) is an

example, in which the personal and involved/interactive registers tend to be narrative as well. This dimension resembles Dimension 1 (personal, involved/stance-focused narration) in web registers (Biber & Kurjian, 2006), showing some similarities between corporate blogs and websites. From a methodological perspective, Dimension 1 in both studies is the first factor which extracts the largest group of co-occurring features in the data, maximizing their shared variance. Accordingly, it is assumed that the corporate blog tend to present information in a narrative discourse with a personal involvement style. It is on the basis of some previous studies that this interpretation is made, i.e., a narrative discourse is characterized by the co-occurrence of past tense and third-person pronouns (Werlich, 1983), and a personal involvement style that is associated with another set of co-occurring features, e.g., *first person pronouns*, *pronoun it* and *indefinite pronouns*, etc. (Chafe, 1982; Biber, 1988). These two functions have been identified as two distinct dimensions in Biber (1988), and co-occur in our CBC. Besides, Dimension 4 (interactive persuasion) of corporate blogs is also a combination of two separate dimensions (Dimension 1: Involved and informational production；Dimension 4: Overt expression of persuasion) identified in Biber (1988). This hybridity showed that the narrative discourse in corporate blogs is not limited to presenting an account of events in the past; rather, it involves people in the events, including people's participation and even some feelings and attitudes towards the events. Similarly, the expression of persuasion in corporate blogs is also accompanied with people's involvement. It is assumed that people's beliefs about, or feelings toward an event or something like purchase intentions for the product plays an important part in narrating and persuading audiences. These mixed functions can be more fully recognised, after a further investigation of the use of some language features from Halliday's (2004) metafunctions, e.g., modal verbs, and pronouns, etc..

The dimensions identified in corporate blogs also exhibit a wide scope of discourse functions that overlap with spoken and written registers (Biber, 1988), e.g., academic prose, press reportage, fiction, letters, conversations, interviews, radio broadcasts, public speeches, as well as web documents (Biber & Kurjian, 2006), e.g., documents sampled from 63 web sites for 'Home' and 1576 documents sampled from 81 web sites for 'Science' ('Home' and 'Science' are two Google categories). Among many previous MD studies, whether by Biber or other

researchers, the most commonly shared dimension is the one involving a narrative function. For example, a narrative function has been identified in Dimension 2 (narrative versus non-narrative discourse) in Biber (1995), Dimension 2 (narrative versus non-narrative discourse) in Biber et al. (2002), Dimension 1 (personal, involved narration) in Biber and Kurjian (2006), Dimension 4 (narrative style) in Grieve et al. (2011), and Dimension 1 (narrative versus expository prose) in Asencion-Delaney and Collentine (2011). Although the discourse function of narration plays an essential role in the production of corporate blogs, it was thought of as having no systematic relationship to spoken and written language. Rather, other functions concerning involved versus informational production, situation-dependent reference or abstract style indicated sharp distinctions between spoken and written registers.

Based on the six dimensions from the factor analysis, six clusters/text types were identified and interpreted, from which two conclusions can be drawn. Firstly, the corporate blog is a hybrid genre, representing a combination of various text types. Comparing the present analysis to previous MD studies on blogs (Grieve et al., 2011) provides two perspectives on the characteristics of corporate blogs. One is that corporate blogs appear to be the standard blog voice, the style of which is highly interactive and conversational. If considering the distinctiveness of corporate blogs in relation to the full range of blog register, corporate blogs display features of a new genre, the variation of which does not fit into any existing text types. Rather, CB text types identified are characterized by a mix of language features and functions co-occurring in one text type. Secondly, diverse text types are used in corporate blogs to achieve different communicative purposes and functional goals. For any corporate blog published by a company, the text types are also varied, which coordinate with each other and work together to achieve their organizational purposes. It could be noted that varied text types in a corporate blog site provide readers a writing repertoire that lays the foundation on various topics and aspects, and the borderlines of the CB text types are relatively fuzzy in comparison with traditional text types.

In the writing repertoire of corporate blogs that shows their hybridity and complexity, the narration characterized by storytelling is a very basic text type, which aims to make sense of

experiences and come to an understanding of events. In the narrative focused texts, some core narrativity features, such as past tense verbs and third person pronouns are frequently used, as suggested by Werlich (1976) and Biber (1988). This can be illustrated in text type 3 (interactive storytelling), a core narrative text type focusing on telling a story. The CB texts from this text type takes up the smallest percentage of corporate blogs; however, this does not impede the narrativity of corporate blogs, since several other text types identified are actually associated with narration, although they focus on 'narrativity' intermediately (Sturgess, 1992). For example, text type 1 (descriptive discourse), text type 2 (technical conversation), text type 4 (evaluative information report), and text type 5 (persuasion) do not primarily focus on telling a story, but include narration as a secondary or simultaneous objective alongside primary objectives and underlying functions, such as scientific exposition, persuasiveness, information presentation, or interpersonal interaction. These text types with the secondary focus on narrative might be called peripheral narrative text types. In these peripheral narrative text types, there are some interactions between narrativity and other factors, including interactivity, and some business or industrial 'objectivity' that influences the use of the core narrativity features. For instance, the use of evaluation features such as the emotional stance verbs, adjectives, and nouns is dominant in text type 4 (evaluative information report). When these stance verbs, adjectives, and nouns are used objectively and followed by inanimate subjects, the focus is shifted from showing a personal viewpoint to some activities or facts. Even in the text type combining both interpersonal involvement and technical information, i.e., text type 2 (technical conversation), a conversational tone is adopted for breaking through minimizing the social distance brought about by the technical information. These examples show that there are some tensions between interpersonal interaction and factual objectivity in a text, or even in a CB text type. Following Biber's (1988, p. 140) claim that instructional writing should be classified as "non-narrative", text type 6 (Instruction) in corporate blogs could be labelled as 'non-narrative' as well, which does not primarily focus on narration as a means for facilitating the understanding of experiences. Rather, texts from this text type are mainly concerned with telling someone what to do and how to do something through an interactive, narrative and somewhat persuasive tone.

Overall, various CB text types constitute corporate blogs as a whole, and they serve different purposes. The text type of interactive storytelling is probably concerned with building positive corporate image, and establishing community between readers and writers; text types of descriptive discourse, technical conversation, evaluative information report, and persuasion are mostly probably associated with developing new product, conducting direct sale to a wide range of audiences; while the text type of instruction may help audiences or customers solve problems by giving instructions and explanation, etc. Among the six text types identified, text type 2 (technical conversation), which takes up the largest percentage of the whole data, focuses on technical issues, and the language employed in this text type functions to close the social distance between readers and writers, by means of 'talking' to their professional rather than amateur readers in a conversational tone. Taken together, the corporate blog includes not only an 'oral-literate' continuum but also a narrativity continuum in which some primary and secondary objectives, such as narrative, interpersonal interaction, persuasiveness, or information presentation, are achieved. Considering that the corporate blog is largely influenced by what industry/business it is from, the relationship between industries that corporate blogs belong to and CB text types was explored and the results of the exploration were further discussed, as presented below.

Firstly, it was observed that most CB texts in text type 1 (descriptive discourse) are from the publishing industry. For blogs from the publishing industry, there has been a shift from traditional paper-based marketing to the use of social media to market directly to potential customers. By using new media, it becomes possible to present readers with more information, such as events related to the book, the newspaper, and more importantly, lead to higher purchasing probability. In spite of this, delivering primarily factual description is still an essential part of marketing in the publishing industry, since this industry is closely related to a descriptive text type. From a linguistic perspective, CB texts from the publishing industry are dominated by language features from text type 1 (descriptive discourse), but possibly lack those features or their communicative purposes of text type 5 (persuasion).

Secondly, it was also observed that CB texts from text type 2 (technical conversation) take the

largest percentage of the whole data and distribute relatively evenly in different industries. In spite of this, among CB texts from text type 2 (technical conversation), most of them are from the industry of information technology, and fewest of them are from the publishing industry. CB texts in text type 2 have less marked language features in general, notably being characterized by a technical and conversational/informal language style. The dominance of CB texts from text type 2 (technical conversation) in the industry of information technology showed that companies from this industry have employed blogs as an effective marketing and communications tool for achieving a fundamental organizational goal: sharing domain expertise with the customers. It is incredibly valuable to share expertise with old customers and bring the best to them; however, considering that some new customers are possibly the layman, rather than experts, a conversational tone with a human voice in language is preferred for establishing customer loyalties. For the firms from the industry of information technology, they do cover broad subjects dealing with technology and some aspects of managing and processing information, such as the use of electronic computers, network administration, server maintenance, and computer software. It is often difficult for the general public to understand these information, due to the use of unfamiliar scientific and technical terminologies, or the complexity of information presentation, etc.. In view of this, the firms try to ease the tension by using some more human language to 'talk' to their customers in a friendly and comfortable environment.

Thirdly, there was a striking contrast between the dominance of some text types in the wholesale and retail trade industry and the industry of e-commerce and online social networking service. Most CB texts from text type 3 (interactive storytelling) and text type 6 (interactive instruction) are from the whole sale and retail trade industry, while most CB texts in text type 5 (persuasion) are from the industry of e-commerce and online social networking service. This seems to suggest that the corporate blogs from wholesale and retail trade industry tend to tell stories, and give instructions rather than persuading their audiences to do something. Traditionally, persuasion is a strategy used for sales promotion in a trading industry, which usually involves selling goods or services directly to consumers; avoiding using it implies another basic strategy of corporate blogs: building loyalty of blog visitors is more

important than selling goods directly to them. In contrast, persuasive texts are found to be more dominant in the industry of e-commerce and online social networking service. This shows that persuasive messages are central to interpersonal influence in online sales and communities, where consumers communicate exclusively via the text.

The exploration of the relationship between CB text types and industries helps establish a link between language characterizing certain text types and the industries that corporate blogs belong to. This investigation suggests that different communicative strategies, e.g., persuasiveness, storytelling, description, etc., are implemented in different corporate blogs, giving rise to different interpersonal influences. With corporate blogs as a platform to bring blog readers and writers together, the informational sources which constantly bring topics for communications plays such an important role in building a long-lasting and trustworthy relationship between customers and the company, thus selling industrial products and creating corporate images. In spite of the importance of interpersonal communications and influences in corporate blogs, some other metafunctions, i.e., field relating to ideational metafunction (what is talked about, and the goals of the text), and mode relating to textual metafunction (the kind of text that is made) from the context of situation, may not be ignored if corporate blogs are mapped onto Halliday's (2004) three metafunctions of language. These situational configurations are inseparable from each other, and what would be focused on are some language features that realize these meanings.

It is decided that a further exploration of ideational, interpersonal and textual functions in corporate blogs should be conducted. This is based on the recognition of a shared interface between SFL's three metafunctions and the principles of credible communication proposed by Maslansky et al. (2010). On the one hand, the three metafunctions of ideational, interpersonal, and textual are the three broad aspects of the meaning system in the process of communication. On the other hand, the principles of credible communication proposed by Maslansky et al. (2010) exhibit certain quality attributes of effective communication. It is particularly the 'Personal Principle' and the 'Plainspoken Principle'[14] that aroused my interests

[14] It has been proposed by Maslansky et al. (2010) that the 'Personal Principle' involves personalizing the message by means of speaking from an individual situation rather than trying to put people in regulations, and the

to explore how personal and plainspoken the CB language is, and what the nature of social activity in corporate blogs is. In particular, the 'Personal Principle' suggests that people should make the communication human by telling personal stories or by using a certain pronoun, such as *we*, *us*, rather than *you*, if they want the audience to connect with a product or an issue. I considered this attribute to be closely related with interpersonal metafunction, and decided to explore the use of pronouns and modal verbs in the current research. Besides, the 'Plainspoken Principle' suggests that people should try to avoid using technical terms, and communicate by using clear and simple language. This attribute gives rise to my decision to explore the lexical density and grammatical complexity from textual metafunction, which are the concrete linguistic implications of mode and may contribute to the readability of a text (Eggins, 2004). Although the textual component is realized in part by such aspects as cohesion, and theme-rheme patterns, the focus will be on the analysis of lexical density and grammatical complexity which tell us about whether the texts have a more spoken or more written character. Another strand of meaning manipulated simultaneously and collectively with interpersonal metafunction and textual metafunction, is ideational meaning, which refers to the on-going social activity. It could not be ignored, because no matter making the communication personal, or plainspoken, or human, all of them help to connect issues to people's experience by means of linguistic resources through which our experience of reality is represented. The language feature selected to represent the ideational function is 'the most common verbs and their process types'.

Overall, on these bases, a further analysis for a better understanding of corporate blogs was then conducted rooted in a combined framework encompassing both Biber's MD analysis and Halliday's SFL theory. Based on the identification of text types in corporate blogs, and the interpretation of them in relation to the industry, an analysis of some salient differences among six text types in terms of ideational, interpersonal and textual metafunctions was conducted (Halliday & Matthiessen, 2004). The language features selected include the most common verbs and their process types, pronouns, modal verbs, lexical density, and grammatical

'Plainspoken Principle', which emphasizes that communicating with clarity and simplicity is a foundation of the language of trust.

complexity. It is expected that this extended analysis is able to provide a more comprehensive view of the functional linguistic variation in corporate blogs.

Chapter 6: Results of an SFL-based analysis of lexico-grammatical features in corporate blogs

6.1 Introduction

The goal of this analysis is to explore the salient linguistic differences among six text types identified in corporate blogs. In particular, the lexico-grammatical features in terms of ideational, interpersonal, and textual functions were explored, by using the framework adapted from Halliday's (2004) metafunctions of language. The language features chosen for analysis are the most common verbs and their process types associated with ideational metafunction, pronouns and modal verbs with regard to interpersonal metafunction, and lexical density and grammatical complexity with regard to textual metafunction. It is assumed that this exploration is able to present a clearer picture of corporate blogs from a functional perspective. It has been revealed in Chapter 5 that corporate blogs represent a rather complex genre embedded in the highly varied context. More specifically, it has been identified that corporate blogs are a hybrid genre consisting of six text types, which serve to achieve different communicative purposes and functional goals, and there is a close relationship between certain text types of CB and particular industries. On this basis, I decided to have a further exploration on how the language features, e.g., pronouns, lexical density and grammatical complexity, that are sensitive to the personal and plainspoken principles in credible communications distinguish CB text types.

Following this introductory section, a brief overview of selected language features associated with Halliday's metafunctions is presented at first (Section 6.2). And then, the results of the analysis of the lexico-grammatical features across six CB text types are presented, which include the most common verbs and their process types in Section 6.3.1, pronouns in Section 6.3.2, modal verbs in Section 6.3.3, lexical density in Section 6.3.4 and grammatical complexity in Section 6.3.5. This is followed by a further discussion of how the six CB text types are distinct from each other in terms of some ideational, interpersonal and textual positioning language features.

6.2 A brief overview of selected language features associated with Halliday's metafunctions

In a general sense, the ideational metafunction is to express ideas or viewpoints through participants (nominal groups), processes (verbal groups), and circumstances (adverbial groups; prepositional phrase) that all relate to the field of the discourse; the interpersonal metafunction concerns social relationships among participants which is mainly expressed by the systems of mood types, modality and person; the textual metafunction concerns organizing language into coherent and meaningful spoken and written texts, which is mainly realized by the thematic structure and patterns of cohesion (Thompson, 2013). In the current research, only certain language features are selected, i.e., the most common verbs and their process types (Section 6.3.1) closely related with field and the ideational metafunction, pronouns (Section 6.3.2) and modal verbs (Section 6.3.3) that are closely associated with tenor and the interpersonal function, lexical density (Section 6.3.4) and grammatical complexity (Section 6.3.5) that are associated with mode and the textual function.

The selection of specific/limited features in the current research is mainly due to the constraints of time and space in completing this study, as well as the nature of the social aspect of CB texts that could be embedded within the different contextual domains of field, tenor, and mode. Furthermore, the social and contextual aspect of CB texts (in terms of field, tenor, and mode) could be reflected in the manipulation of some principles of credible communication proposed by Maslansky et al. (2010). Among the four principles of credible communication, i.e., 'be personal', 'be plainspoken', 'be positive' and 'be plausible', what have been linked to the SFL's three metafunctions are the personal principle and the plainspoken principle. In terms of the personal principle, it consists of some components, such as 'make it human', which suggests using personal language, and telling human stories, thus helping to connect issues to people's current experiences, and 'make yourself real', which means companies should put a human face on a corporate act and build trust by using 'we' and 'our' rather than 'you' or 'me'. As for the plainspoken principle, it is suggested that plain and simple language is critical to building trust. Presented below is an elaboration on how the selection of

language features from three metafunctions are triggered by the principles of credible communication proposed by Maslansky et al. (2010).

Firstly, the selection of 'the most common verbs and their process types' from the transitivity system of the field of discourse is triggered by the component of 'make it human' in the personal principle. As suggested in Maslansky et al. (2010), in order to close distance between the company and their audiences, it is better to tell relevant and personal stories by using human language, rather than merely presenting some factors or statistics. In the world of narrative, a universally accepted assumption is that a narrative text organizes time and tells a story, which might concerns a personal or a social life. All these correspond with the transitivity system that encompasses the world of a person's experience in sensing and being, etc. In the linguistic theory of narrative structures, a 'doing versus saying' distinction is at the centre (Franzosi, 2010), although not all sentences of a narrative are about action, doing or saying. And even by large, "non-story elements may be found in a narrative text, just as story elements may be found in a non-narrative text. A novel may well include the description of a cathedral, and the description of a cathedral, say in a guide book, may include the story of its construction" (Rimmon-Kenan, 1983, p. 16). In view of this, it is worthwhile investigating the use of process types of the most common verbs across a range of CB text types. In the current research, the process type from the system of Transitivity is the centre of attention, and the research in Circumstances is left out. There are several types of circumstantial elements in Halliday and Matthiessen's (2004) model, including time, place, degree, reason, etc. The place and time adverbials have been counted as a variable in the MD analysis, and it will not be further explored in this extended study.

Secondly, as for the selection of pronouns and modal verbs from the tenor of discourse, it is triggered by the component of 'make yourself real' in the personal principle, which suggests that the use of particular pronouns, such as 'we' and 'us' rather than 'you' or 'me', is able to put a human face on a company, as well as put the company on the same side of the audience, thus making the company real, and building trust on the reader. In the current research, not only the first person pronouns, such as 'we', 'us', 'you' and 'me', but more pronoun types

including first, second and third person pronouns will be explored. This exploration can help reveal the writers' perceptions of their own role in the communication, as well as their relationship with expected readers. Also related to the pronouns in the interpersonal metafunction is modality, which enables people to convey varying degrees of the probability or likelihood (possible, probably, certain), the usuality (sometimes, usually, always) of something being, obligation (must, should, may) to do what is demanded, and inclination (willing, want to, determined) to do what is offered (Halliday & Matthiessen, 2004). By this means, the speaker or writer shows their opinions, or their attitudes towards a proposition, indicating an association between modality and viewpoints. In the model of viewpoints realized by modality, two basic concepts are particularly important: internal and external points of view (Simpson, 1993). The internal point of view suggests that the story is mediated by the first-person narrator's view, and often represents a subjective fixed point of view. This viewing stance is characterized by the use of first-person pronouns (I, we) and verbs expressing feelings, thoughts, and perceptions, etc. Besides, the external point of view represents an objective, and neutral stance of participants involved in the story. In the current analysis, I do not consider the conventional classification of modality (moralisation and modulation) for encoding both verbal modality and modal adjuncts at varying degrees of polarity. Rather, I focus on a component part from the modality system, i.e., the use of modal auxiliaries. In particular, Biber's (1988, p. 241) classification is adopted, which includes possibility modals (can, may, might, could), necessity modals (ought, should, must), predictive modals (will, would, shall). This exploration is not only for simplicity but for centring on the language features of first-person pronouns and verbs in the internal and external points of view in the model of viewpoints realized by modality (Simpson, 1993).

Thirdly, related to textual metafunction, both of lexical density and grammatical complexity are the concrete linguistic implications of mode, and are chosen for the empirical analysis of the CBC. They are among the factors that may contribute to the complexity of a text, and may influence the linguistic differences between the extremes of the spoken and the written language (Eggins, 2004). As stated by Martin (1992), in the discourse of mode, spoken language is concerned with the process, language in action and, therefore, it is less complex.

Written mode of language, on the other hand, is related to the product, i.e. language in reflection that makes this type of language more abstract and as a result more complex. In contrast, Halliday (1985) claimed that spoken language is considered grammatically complex and written language is lexically dense. The analysis of these two language features is, in some degree, inspired by the plainspoken principle in Maslansky et al. (2010): the use of clear and simple language plays an important part in building trust and promoting interpersonal communication in business world. If comparing those principles of credible communication by Maslansky et al. (2010), it could be found that they are highly interrelated, and may not be approached in a linear way. For example, the plainspoken principle and the component of 'make it relevant' in the personal principle is similar to each other. Both of them basically follow the rules of simplifying language as well as 'talking to the audiences at their level'. These two points seem to be raised from a 'not strictly logical mind', since it does not take into account the great diversity of target audiences, and thus it tends to be inappropriate to communicate with some sophisticated customers by using simple language, and vice versa. Whether or not this paradox can be resolved, and how it is resolved are full of uncertainty, since it is not sure how language can be simplified when sophisticated terminology is required in a business-to-business audience communication in many industries. As claimed in Maslansky et al. (2010), in business communications, the amount of knowledge that the audience has is often overestimated, and the sophisticated terms have been over-used. It is still uncertain whether or not this situation is the same in corporate blogs; therefore, the use of lexical density and grammatical complexity across six text types was investigated. Since the fundamental goal is to reveal the complexity of CB language, some other language features, e.g., theme-rheme patterns, cohesion, are not investigated. Particular patterns of cohesion[15] are significant semantic predictors of the contextual variable mode. Reference is of particular interest concerning the spoken and written registers distinction: it has been claimed that particularly the distinction between exophoric and endophoric reference (Halliday and Hasan, 1976) should shed light on the spoken-written distinction. A corpus-based study on a certain cohesive device is applicable; however, due to the big size of CBC, which results in a very

[15] cohesive markers—grammatical (reference, ellipsis, substitution) or lexical markers (such as repetitions, superordinates, hyponyms and antonyms) (Halliday and Hasan, 1976)

complex process in analysis, and the space constraints, it could be taken into account in the future study. Besides, the study of theme-rheme patterns could be conducted from a corpus-based perspective; however, they have been mainly for the identification of coherence, i.e., textual connectivity, of the text (Feifei, 2004), or the genre variation (Francis, 1989; Nwogu, 1989). It is believed that the same genre represent a similar contextual configuration, that is, they show common characteristics in terms of field, tenor, and mode of discourse, and the field, tenor, and mode can be realized in topical, interpersonal, and textual themes respectively (Halliday and Hasan, 1976). It seems that the studies on theme-rheme patterns have a relatively loose connection with the spoken/written distinction and the complexity of a language; therefore, the language feature of theme-rheme patterns is not adopted in the current research either. What follows are the results of the exploration on how the language features selected from three metafunctions in SFL differ across different CB text types.

6.3 Results

6.3.1 The most common verbs and their process types

The investigation of the most common verbs and their process types across six text types showed that there is a significant difference in their distributions. As shown in Table 6.1, the verb ***using*** is most frequently used in text type 2 (technical conversation) and text type 4 (evaluative information report). Besides, the verb ***said***, ***found***, ***need*** and ***want*** are the most common verbs in text type 1 (descriptive discourse), text type 3 (interactive story telling), text type 5 (persuasion) and text type 6 (interactive instruction) respectively.

Table 6.1: The most common verb in each text type (equalized to the occurrences per 100000 Words)

Text types	Category	Verb forms	Occurrences	Process types
Text type 1	VVD	said	135	verbal
Text type 2	VVG	using	72	material
Text type 3	VVD	found	225	mental
Text type 4	VVG	using	68	material
Text type 5	VVO	need	125	mental
Text type 6	VVO	want	101	mental

Besides, the most common verbs in the CBC are also explored. Presented below are the top fifteen frequently occurring verbs in the CBC (Table 6.2), among which ***using*** is most frequently used, and the verb ***make*** is the only one that takes all the tenses and patterns (***make, makes, made, has/have made, making***), constituting a great proportion of occurrences.

Table 6.2: The fifteen most common verbs in the CBC

Rank	Category	Verb forms	Occurrences	Process types
1	VVG	using	419	material
2	VV0	need	383	mental
3	VV0	make	295	material
4	VVD	said	289	verbal
5	VV0	know	278	mental
6	VVG	making	238	material
7	VV0	want	212	mental
8	VVN	used	192	material
9	VVZ	makes	189	material
10	VVG	working	181	material
11	VVZ	comes	166	material
12	VVD	made	164	material
13	VVZ	provides	149	material
14	VVN	seen	136	mental
15	VVN	made	129	material

The dominant use of ***using*** and various forms of the verb ***make*** indicates the focus on particular activities in corporate blogs. Semantically, they belong to the activity verb, serving as process in different material clause types (Biber et al., 1999). The verb ***using*** co-occurs with not only animate but inanimate subjects, and is restricted to a limited range of meanings, i.e., centring around the customer usages of products and services, as can be seen in the following examples.

6.1) So, wireless carriers in many instances have contracts in place for existing backhaul facilities that they are ***using*** to handle the traffic that their customers are generating. (cluster 2, Verizon20120329.txt)

6.2) More people are ***using*** location-based mobile services primarily due to consumer incentives offered when people check in. (cluster 4, Corporateeye56.txt)

6.3) Today, we'd like to share some updates we've made to the LinkedIn Developer Platform to continue fostering creative innovations ***using*** LinkedIn's technology. (cluster 2, LinkedIn20120807.txt)

6.4) The CloudFront location in Sydney supports the entire array of CloudFront features2

including support for dynamic content, low minimum content expiration periods, live streaming to multiple devices ***using*** FMS2 4.5 or smooth streaming, streaming media, private content, invalidation, and custom origins. (cluster 4, Amazom20120619.txt)

Sample 6.1 is a text excerpt from the file named 'Verizon20120329.txt', which is from **Verizon-PolicyBlog**, a blog for Verizon Communications Inc. This post starts with CEO James Crowe's suggestion that the solution to the wireless spectrum shortage is bringing more wires, rather than spectrum to market. In sample 6.1, ***using*** collocates with 'facilities', and the whole sentence concerns the use of facilities for handling the traffic that customers generate in wireless companies. In sample 6.2 and sample 6.3, the verb ***using*** collocates with 'services' and 'technologies' respectively. In all of these samples, as well as other samples in CBC, the verb ***using*** is basically an activity for sharing some information, including services, skills and knowledge, etc., with the customers. It is a way for companies to clearly communicate their brand's philosophy with the customers, and to encourage trials of a new product, or services, thus building brand loyalty as well as a long-term relationship with the customers.

Another frequently occurring verb ***make*** not only refers to physical activities, but also being commonly causative in meaning. Among the total occurrences of the verb ***make*** in CBC, a relatively smaller number of them occur with animate subjects functioning as a concrete entity to the verb (38 out of 295), as can be seen in the following examples.

6.5) But the rewards are similar: in the office, we ***make*** a connection with a reporter and form a relationship;… (Cluster 2, UnboxedThoughts20120802.txt)

6.6) I'd been at it for almost two months and between marketing materials, website design and hiring an assistant, I'd yet to actually get out there and ***make*** a sale. (Cluster 2, LinkedIn20120913.txt)

In contrast, the verb ***make*** is more commonly used in a causative sense with inanimate subjects (257 out of 295), i.e., its syntactic use is typical exemplified causative construction in which the 'make' verb serves as a causative index. The following examples illustrate this:

6.7) That's awesome! It'll ***make*** life so much easier!! (cluster 1, rub17.txt)

6.8) It's hard enough to type in payment info via the web on a small screen computer or tablet, but the smaller screen of mobile devices ***make*** this process significantly more tedious.

(cluster 1, 1.txt)

The verb ***make*** is also the only lexical verb taking almost all tenses and patterns (VV0, VVD, VVG, VVN, and VVZ). The frequencies of these tenses and patterns are varied: VV0 (make, 234), VVD (made, 129), VVG (making, 239), VVN (made, 164), and VVZ (makes, 189). These frequencies show a basic difference in verb ***make*** used in the five tenses. There is not much functional implication from these results; rather, the use of different forms of the word ***make*** could be further categorized into different patterns in accordance with Biber's work (Table 6.3).

Table 6.3: Verb patterns (Biber et al., 1999, p. 384)

SV	intransitive pattern
SVA	copular pattern with obligatory adverbial
SV+ A	intransitive pattem with optional adverbial
SVOd	monotransitive pattern with noun phrase as object
SVPs	copular pattern, with adjective phrase or noun phrase as the subject predicative element
SVOiOd	ditransitive pattem, with both indirect and direct objects
SVOdPo	complex transitive pattern, with adjective phrase or noun phrase as the object predicative
SV + complement clause	pattern with a complement clause following the verb
SVO + complement clause	transitive pattern with an object and a complement clause following the verb

Presented below are the sample sentences illustrating the patterns of various forms of the verb ***make***.

- SVA (with obligatory adverbial):

6.9) Forget for a moment what you have to disclose because of regulations and quarterly filings and think instead about other things that ***make*** up your company. (cluster 1, 59.txt)

6.10) Both websites are bidding on their own brand terms, but branded paid search only ***makes*** up about 25 percent of YouTube's SEM traffic, as opposed to Hulu, where over 50 percent of their SEM traffic is branded. (cluster 1, comscore20120430.txt)

- SVPs (with an adjective as predicative):

6.11) There are several reasons for this, chief of which is that investors will view the greater transparency into operations and management thinking about the future as enabling them to ***make*** better estimates about future cash flows.

- Monotransitive SVOd (with noun phrase as direct object):

6.12) They ***made*** some very interesting points during the webinar which resonated with what we've been sharing on this blog. (cluster 2, UpsideLearning20120724.txt)

6.13) She ***makes*** a personal investment in all of her fellow associates keeping the lines of communication open and making time to visit every department to offer encouragement or just listen. (cluster 2, Marriott20120509.txt)

- Ditransitive SVOiOd (NP):

6.14) So, I have an idea to share that will allow you to still include confetti----***make*** it part of the design element even---but keeps the mess at bay. (cluster 2, 116.txt)

- Ditransitive SVOdOi (with indirect object in a prepositional phrase):

6.15) I think for us as market researchers, who need to ***make*** recommendations to our clients about whether to launch a new product into the market, this is a very important lesson. (cluster 2, rs0621.txt)

6.16) While most days at Janssen we strive to ***make*** a difference for patients through our research to create simpler, safer and more effective treatments for hepatitis C, on August 2, at the second annual White House World Hepatitis Day Event, I had the opportunity to share my thoughts about how Janssen can help make a difference in another way: through public and industry partnerships. (cluster 2, JNJBTW20120907.txt)

- Ditransitive with only an indirect object SVOi:

6.17) They also visited local attractions and ate their first Chesapeake Bay crab. So next time you open your refrigerator and cupboards and see a mish-mash of ingredients, be creative and try to ***make*** a meal. (cluster 2, Marriott20120723.txt)

6.18) Frozen and canned vegetables are a great option. We wouldn't have ***made*** it, frankly, without them. (cluster 3, WholeFoods20120222-2.txt)

6.19) Customers are ***making*** choices. They are changing carriers and they clearly are making decisions to go in a new direction if they aren't satisfied with what they have.

(cluster 4, Verizon20120329-2.txt)

- Complex transitive SVOdPo (with an adjective as object predicative):

6.20) It ***makes*** me crazy to find out that many people still do not know that Xerox has Online support for our full line of printers and multifunction devices. It is an amazing resource that is a simple click away. (cluster 5, Xerox20120907-4.txt)

6.21) My mind drifts from the scary face and expands to my surroundings. Besides us, surely no humans have ever had this specific view - only the condors. Wow, I think, how cool. But it also ***makes*** me a little sad. (cluster 5, patagonia20120413.txt)

- Complex transitive SVOd + complement clause:

6.22) We realize this is not an ideal experience and that tweaks need to be made to the new search. So for now, every customer on delta.com is getting the same search engine, and we're ***making*** changes to the new search functionality to offer more search results and to provide customers with the right balance between quality and price in those results. (cluster 6, Delta20120518.txt)

Among the most common verbs from different text types, ***said*** is the most frequently used verb in text type 1 (descriptive discourse). It is the only one that could be categorized into verbal process, which is a process of asking, telling, commanding, showing, or offering, etc. The dominant use of ***said*** is to report the speech by a person him/herself or people in a past utterance, as shown in the following examples. This result is close to Biber et al.'s (1999) report on the LSWE corpus: the verb ***say*** is the most common verb overall in the LSWE Corpus and the only verb to be extremely common in more than one register.

6.23) As Nixon ***said*** in the handwritten letter he sent thanking me for the return of the purloined pieces, it was the high point of his presidency, and he hoped, the accomplishment for which he would be remembered after the scandal of Watergate dimmed. (cluster 1, nytimes1-20120221.txt)

6.24) Water was up to 4 feet deep in many areas of Hancock County as the storm system lumbered off the mouth of the Mississippi River, Brian Adam, the county emergency management director ***said***. (cluster 1, NYT165.txt)

The close relationship between the verb ***said*** and text type 1 (descriptive discourse) shows that text types 1 might be very close to spoken language or as being set in a dialogical passage. The following example shows this.

6.25) "By partnering with eBay, passengers can now shop a wide selection of listings on the site from 35,000 feet, "said Ash ElDifrawi, Gogo's chief marketing officer.(cluster 1, IAG20120814.txt)

Semantically, the verb ***said***, as a verbal process, could be further categorized into different functions, including informing (e.g., 6.26), stating (e.g., 6.27, 6.28), suggesting (e.g., 6.29), telling and saying (e.g., 6.30).

6.26) After briefly making landfall, Hurricane Isaac has once again moved over water, the National Hurricane Center ***said***. (cluster 1, 167.txt)

6.27) Mark Taylor, the chief operating officer of In-N-Out Burger, ***said*** in an e-mailed statement that the slaughterhouse provided the company with beef in chuck form that it used to make its own patties. (cluster 1, 173.txt)

6.28) Raw footage was handed over to United States Department of Agriculture officials last Friday. Its Food Safety and Inspection Service ***said*** in a statement this week that it had received disturbing evidence of inhumane treatment of cattle at the meat company and suspended the work of inspectors there on Aug. 19, effectively forcing the plant to halt its slaughtering operations. (cluster 1, NYT173.txt)

6.29) However, comScore data shows that while search is being leveraged by the major premium video channels like Hulu, the same cannot be ***said*** of Web TV. (cluster 1, comscore20120430.txt)

6.30) They ***said***, We can't do anything about it. (cluster 1, nytimes20-20120125.txt)

Overall, in all the sample sentences above, the central participant is the sayer, who is engaged in a human or human like communication. In text type 1 of CBC, the vast majority of the use of the verb ***said*** (62 out of 74) could be categorized into the function of saying and telling, with an animate subject. In contrast, much fewer sentences encompassing the verb ***said*** (12 out of 74) are followed by inanimate subject, indicating a certain situation. Besides, the past tense is not always confined to past time reference, but referring to present or the future time. Examples include:

6.31) At the 6 a.m. briefing at corps district headquarters, Chris Accardo, chief of operations for the district, ***said*** that gauges at the canal show the water level has risen to levels as high as six-and-a half feet C at or beyond the margin of safety for the floodwalls. (cluster 1, NYT165.txt)

6.32) I've had nightmares, but nothing like what you went through, and many Argentines. And in our case, we were not heroes, we had lost a war. Our commanders ***said*** people would throw eggs and shout at us. (cluster 1, guardian14-20120302.txt)

6.33) Years ago after improvements were announced for the DI technology some experts ***said*** that DI technology would become the dominant form of offset technology. (cluster 1, Xerox20120911.txt)

In text type 5 (persuasion), the most common verb is ***need,*** the occurrence of which is 125 (equalized to the occurrences per 100000 Words) with a tag of VV0, and its occurrence before equalization is 114. It is also the second most common verb among the verbs tagged with VV0, VVD, VVG, VVN, and VVZ in CBC. Among the total occurrences of the verb ***need***, 31 of them can be categorized into mental process characterized by a consciousness of the Sensor who wants something or put some opinions on other people, as can be seen in the following examples:

6.34) You ***need*** additional steps for Eclipse Juno, because it does automatically detect and download MTJ (Mobile Tools for Java) plugin. The following instruction is for Eclipse Indigo. (cluster 5, ora20120925-20.txt)

6.35) All you ***need*** is a relatively small space for the Espresso Book machine and you can print and sell books on demand. (cluster 5, Xerox20120911-2.txt)

Besides, 83 out of 114 ***need*** is within the structure of ***need to*** which is marginal modal structure explaining ideas like *necessity*, or *obligation.* As a type of modal verb, it can be semantically categorized into the 'intrinsic modality' (e.g., 6.36-6.39) which speaks about agent-controlled actions and events, rather than the extrinsic one which relates to the reality or the logical status of events (Biber et al., 1999).

6.36) You can't mark without a scheme. You ***must*** make a scheme. (Biber et al., 1999, p. 485).

6.37) We ***shall*** not attempt a detailed account of linguistic categories in this book,… (Biber et al., 1999, p. 485).

6.38) You can read the full tutorial here: http: Of course, if you use this technique for my challenge, you still ***need to*** think of an interesting and fun way to use some Fiskars tools. (cluster 5, Fiskateers101.txt)

6.39) Once I stopped laughing and read further, the point he was making was that often we already have much of what we ***need*** to deliver an exemplary customer experience. (cluster 5, AVAYA20120625.txt)

Both of ***need*** and ***need to*** from text type 5 (persuasion) share something in common: almost all them co-occur with the second person pronoun ***you***, thus serving as a very interactive and persuasive language device to provide their opinions on issues intrinsically, give judgements about a topic, or give emphasis to an argument, etc.. All these work together for the achievement of an 'overall purpose' of expositions, discussions, or advertising, etc. in text type 5 (persuasion).

6.3.2 Pronouns across six text types

This section presents both quantitative and qualitative analysis of various personal pronouns based on their frequency across six clusters/text types. Frequency counts were employed for the identification of the differences of first-person, second-person, and third-person pronouns across six text types. Moreover, the discourse functions of those pronouns were interpreted qualitatively on the basis of actual occurrences in the sample texts and the context. Prior to the frequency counts, six directory files were ready, each of which included texts from six CB text types respectively. The number of texts included in each file was varied: the file named text type 1 includes 120 texts, the file named text type 2 includes 525 texts, the file named text type 3 includes 4 texts, the file named text type 4 includes 160 texts, the file named text type 5 includes 193 texts, and the file named text type 6 includes 18 texts. After the data was processed in the software AntConc (http://www.laurenceanthony.net/software/antconc/), the frequency of each kind of personal pronouns was counted, and the results were presented**.** For example, in the main window of AntConc, the frequency of the personal pronoun 'she' in text type 1 is able to be counted directly, and the word tokens of this text type is also presented, as shown in Figure 6.1. And the frequency count of each kind of personal pronouns in each text type is presented in Table 6.4.

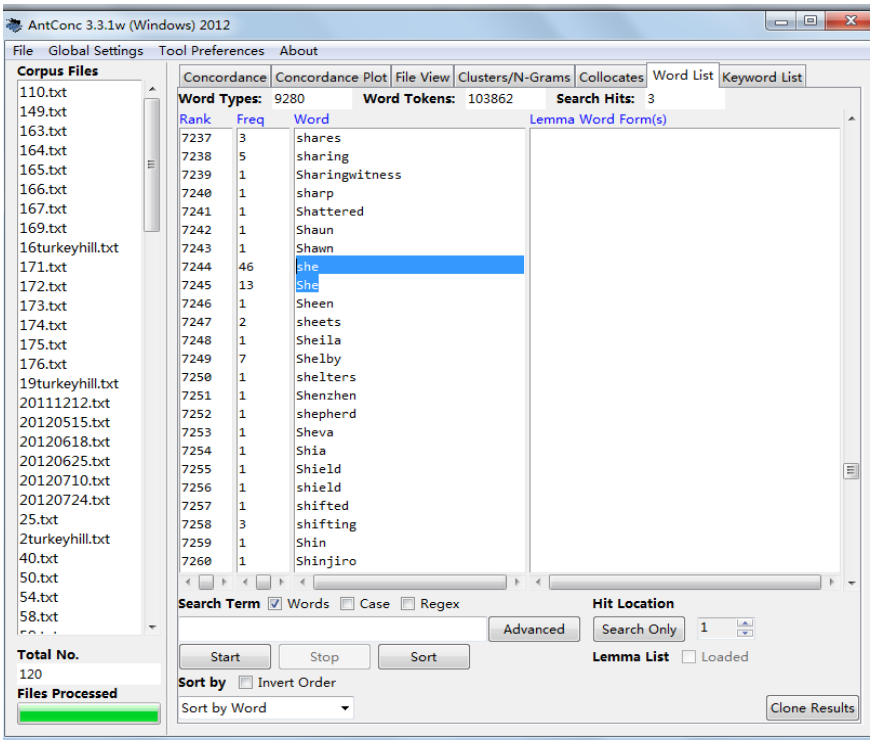


Figure 6.1: Frequency count of ‘she’ in Text type 1

Table 6.4: Frequency count of each personal pronoun in six text types

	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6
	descriptive discourse	technical conversation	interactive story telling	evaluative information report	persuasion	interactive instruction
I	480	2137	91	470	730	49
we	292	1883	50	444	421	131
you	221	2061	35	550	2548	126
they	140	868	12	300	333	33
he	108	290	9	155	59	1
she	59	281	7	110	65	1
them	48	367	5	105	220	28
her	61	219	4	100	80	4
him	12	45	2	22	9	0
me	64	427	23	68	119	2
us	52	310	10	79	74	19
it	345	1989	29	643	1122	61
Total Number of Words	103862	284047	3137	91770	96847	11000

The data would be more plausible if they are normalized to the occurrences per 10000 Words, as shown in Figure 6.2.

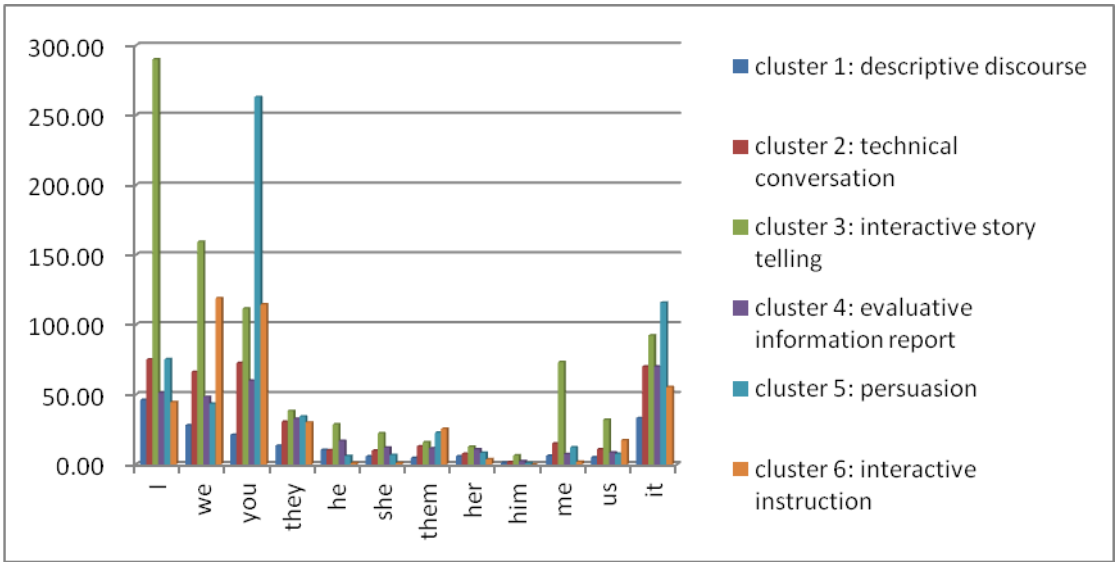


Figure 6.2: Frequency of pronouns across clusters (Equalized to the Occurrences per 10000 Words)

In this figure, the difference in frequency among the categories of first-person, second-person,

and third-person pronouns across six text types is substantial. First-person pronoun ***I,*** first-person plural pronoun ***we***, and second-person pronoun ***you*** occur far more frequently than other types of personal pronouns in all six text types. Among these most frequently used pronouns, the occurrences of ***I*** and ***we*** in text type 3 (interactive story telling) exceed considerably those in other text types, the occurrence of ***you*** in text type 5 (persuasion) exceeds considerably that in other text types, and the occurrences of ***you*** and ***we*** in text type 6 (interactive instruction) are smaller than those in text type 3 and text type 5, but bigger than those in other text types.

Previous researchers have commented that personal pronouns, such as ***you, I,*** and ***we,*** place the speaker and the addressee within a specific social context (Kashima & Kashima, 1998). That is, ***I*** emphasizes the social distance between the speaker and the social context (Kühnen, Hannover, & Schubert, 2001), but ***we*** and ***you*** index social interdependency between the speaker and addressee: ***we*** emphasizes shared identity (Simmons, Gordon, & Chambless, 2005), inclusiveness or exclusiveness of the parties (Biber, 1995; Kuo, 1999; Pennycook, 1994; Rounds, 1985; Spiegelberg, 1973); ***you*** reflects social distance between the parties (Kern, Lee, Aytug, & Brett, 2012). What follows is the elaboration of the use of ***I*** in text type 3, which shows several categories as labelled.

1) ***'I'*** as existential subject

This role can be signalled by relational processes that are concerned with the process of being in the world of abstract relations. The types of relational process are quite complex and controversial (Eggins, 2004). In the current analysis, it is sufficient to identify the following two principal process types: the attributive mode and the identifying mode in intensive processes. As can be seen in the following examples, ***was*** in 6.40, and ***am*** in the sentence 'I ***am*** a First Officer for Delta Air Lines' are the identifying mode, and the use of ***was*** in 6.41 and ***am*** in the sentence 'I am based in the New York area' are the attributing mode.

6.40) As an African-American female, I **was** a double minority and for four years I **was** the only girl in many of my courses.(cluster 3, Delta20120323.txt)

6.41) My name is Dawn Cook and I **am** a First Officer for Delta Air Lines. Currently in my fifth year with Delta, I fly the Boeing 737 (800/700). I **am** based in the New York area, so I fly out of JFK, LaGuardia and Newark a lot. (cluster 3, Delta20120323.txt)

6.42) I **was** completely surprised as neither of us had hinted during our hours of

conversation that we had an attraction to one another. I felt butterflies consume my body, I **was** smitten; I smiled and ran onto my almost departing plane. From that day forward, not a day would go by without talking with one another. (cluster 3, Delta20120419.txt)

2) ***'I'*** as the recounter of the story

This particular role is often signalled by the pairing of first person pronouns with material process verbs (i.e. "doing" verbs like work, read, interview, collect) (Halliday & Matthiessen, 2004) used in the past tense. Some sample sentences are presented below:

6.43) Before we had reached our cruising altitude, I **found** myself engaged in a pleasant conversation with this surprisingly charming neighbour. (cluster 3, Delta20120419.txt)

6.44) Inside of the handbag, I **found** two blank Delta boarding passes that Chris had written on. (cluster 3, Delta20120419.txt)

6.45) I **entered** a house full of people and food, including eight different kinds of pie and the first turducken I had ever seen. (cluster 3, WholeFoods20111114.txt)

6.46) I **made** two loaves of bread every week for a mere $1.25. It makes sense: if you do the labor, you keep the savings. (cluster 3, WholeFoods20120222-2.txt)

3) ***'I'*** as the opinion-holder

In this category, the first person pronoun ***I*** can be taken as an indication of authorial positioning or stance by writers (cf. Hyland, 2002).This role, when it involves the co-occurrence of the first person pronoun and verbs, describes what Halliday and Matthiessen (2004) termed mental processes of cognition. The following sample sentences show this:

6.47) I **hoped** that, with any luck, I would get one of the opportunities and start taking a step in the right direction for my career. (cluster 3, Delta20120323.txt)

6.48) What I **thought** was a purse, ended up being a distraction to hide the ring. (cluster 3, Delta20120419.txt)

A relatively dominant use of the first person pronoun ***I*** in text type 3 (interactive story telling) indicates a speaker/writer oriented communication in which the speech is more subjective, and the speakers intend to narrate a story. The first person pronoun ***I*** has also been used to create a 'relationship' with the audience, because the use of ***I*** is able to show commitment to the audience and personal involvement in issues, thus giving the speaker a personal voice that distances him/her from others. This is the advantage of using ***I***; examples of personal qualities of a blog writer include not only some personal routine work, but some ideas or moral principles on something. It is assumed that the use of ***I*** is able to build trust and credibility in communication.

The pronoun ***we*** is frequently, yet somewhat differently, used in text type 3 and text type 6, in terms of the exclusive-inclusive distinction. Many researchers have already distinguished between "exclusive ***we***" and "inclusive ***we***" (Biber, 1995; Kuo, 1999; Pennycook, 1994; Rounds, 1985; Spiegelberg, 1973). 'Inclusive-***we***' includes both the writer and the target audiences/readers, whereas 'exclusive-***we***' only refers to the writer and the other people associated with the writer. The CB texts from both text types are single-authored, while the use of ***we*** may indicate an inclusion of others in the utterance. Among the text samples from text type 3, '***we***' refers to the blog writer and his/her companions in a story, without including the readers, as presented below.

6.49) After arriving in Atlanta, **we** found ourselves wandering around and looking for somewhere to grab some food. It was at that time **we** realized that we did not even know each other's name, as **we** had not formally introduced ourselves to one another. (cluster 3, Delta20120419.txt)

6.50) After dinner **we** went for a walk around the neighbourhood to enjoy the beautiful autumn weather. Back at the house, a co-worker got out his guitar and **we** lounged around the living room listening to music and talking till the sun went down. (cluster 3, WholeFoods20111114.txt)

I suppose the pronoun ***we*** in text type 3 (interactive storytelling) may not necessarily be involved in the issue of inclusive/exclusive distinction; rather, it signals an access to a wider scope, as well as an engagement with narrative texts in an unusual way. That is, the pronoun ***we*** creates an environment, in which the narrator and the reader are coordinating with each other, or that the narrator plays the role of a tourist guide to the reader. At the same time, the description of the story is like a film, and the reader can naturally imagine what happened. In a narrative space where an inclusive ***we*** is used, the narration would possibly blend the writer's as well as reader's experiences.

Similarly, text type 6 (interactive instruction) is dominated by *exclusive-**we***. As can be seen in Figure 6.3, many concordance lines from AntConc showing the use of '***we***' in text type 6 are presented, which includes a huge number of examples in varied contexts and situations.

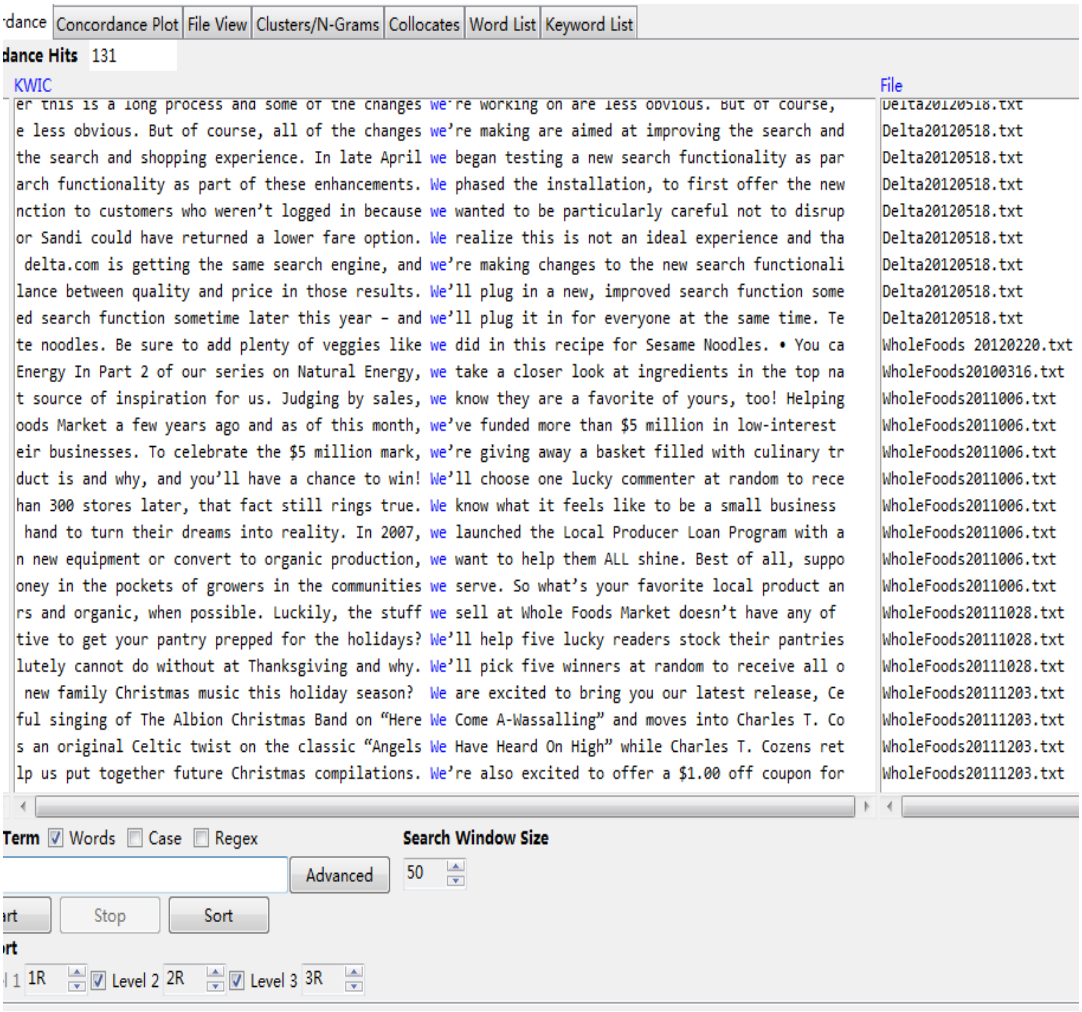


Figure 6.3: Concordance lines of the use of 'we' in text type 6

In all the occurrences, a wide range of discourse functions could be identified, such as explaining what to be done or what has been done, expressing wish or expectation, and expressing gratitude, as illustrated in the following examples.

6.51) **We**, like all international airlines, are required to comply with all applicable laws governing entry into every country **we** serve. (cluster 6, Delta20110623.txt)

6.52) It was so good **we** now plan to ring in Christmas every year with Chinese barbecue. (cluster 6, WholeFoods20111212.txt)

6.53) **We** have a lot of exciting things yet to come and will continue to deliver program improvements based on our members' feedback. (cluster 6, Delta20100729.txt)

6.54) **We** hope these changes to our policies reflect the true respect we hold for our service men and women and again demonstrate our appreciation as both a company and as individuals who benefit from the freedom our troops defend. **We** also hope this echoes our commitment to listening, learning, and our continued dedication to constant improvement. (cluster 6, Delta20110607.txt)

A closer look at the sample sentences using '***we***' from this text type seems to reveal a contradiction between the fact that all CB texts from this text type are single-authored, and a

relatively more dominant use of ***we*** than ***I***. The goal of using ***we*** in text type 6 (interactive instruction) is to make the readers believe that the decisions or any instructions that are being made are the right ones. Especially among those sentences using exclusive ***we***, the main purpose is to share responsibility, that is, actions are not only the responsibility of one individual. Although most of ***we*** is not categorized into an inclusive-***we***, it does not hinder the establishment of solidarity between the writer and readers, because using ***we*** rather than ***I*** may indicate the exclusion of some other groups, which implies a division between ***us*** and ***them***. Simultaneously, a sense of togetherness can also be created, owing to some interactive 'oral-like' expressions, such as the occurrence of ***you*** and ***we***, which promotes the integrity of the reader-writer community (e.g., 6.55, 6.56).

6.55) Now through March 31st, **we** hope that **you** will join fellow shoppers, team members, Supplier Alliance for Microcredit vendor partners and online donors in raising $4.1 million to help impoverished people living in communities around the globe where Whole Foods Market sources products. …**We** thank **you** for your support! (cluster 6, WholeFoods20120222.txt)

6.56) **We** thank **you** for your questions and remain happy to answer any additional questions you may have. **We**'ll also be sharing updates along the way to keep **you** informed. (cluster 6, Delta20110607.txt)

In the text types other than text type 3 and text type 6, it is also found that *exclusive-**we*** is much more frequently used than *inclusive-**we*** (e.g., Figure 6.4, Figure 6.5). No matter whether it is a descriptive discourse or a technical conversation, all of them are in an interpersonal situation activated by a certain degree of authorial presence. Under most circumstances (1009 out of 1020 CB texts), ***we*** is used in single-authored texts instead of ***I***. In some CB texts (those from blog No. 39, 63, 73, 80, 89, 267, 271, 276, 284, 530, 532), ***we*** is used in texts written by more than one author. These two options could be situated on a continuum. The decision to use ***we*** by writers of single-authored texts may indicate an intention to reduce personal attribution, while the dominant use of exclusive-***we*** may impede the establishment of solidarity and the reducing of social distance. As claimed by Pennycook (1994, p. 174), "there is an instant claiming of authority and communality in the use of ***we***". Hyland (2001) also argues that the use of ***we***, as a plural meaning, seems to have a temporary dominance by giving the writer the right to speak with authority, it is basically a strategy, by means of which the writer can simultaneously reduce their personal intrusion and yet emphasise the importance that should

be given to their claims. This situation also applies to corporate blogs, which is on a dynamic continuum between sharing some information and reducing the personal attribution.

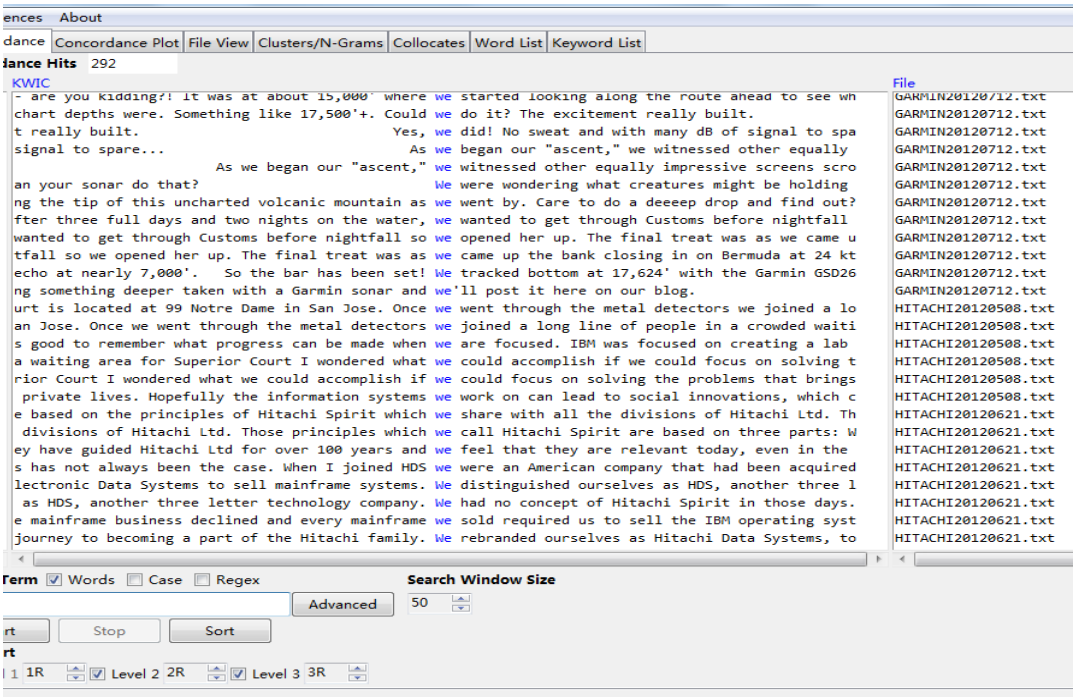


Figure 6.4: Concordance lines of the use of 'we' in text type 1

Figure 6.5: Concordance lines of the use of 'we' in text type 2

When it comes to ***you***, it is most frequently used in text type 5 (persuasion), rather than other text types. This result shows that the pronoun ***you*** plays a prominent role in a persuasive

discourse, in which a bond between the persuader and the consumer/reader could be established. It is claimed by Smith (1985) that ***you*** is the most interactive form among personal pronouns, because it explicitly acknowledges the presence of the reader and addresses to them directly. It is a marker of 'one-to-one' dialogue, although ***you*** could be used to refer to a specific person as well as people in general. In order to determine whether '***you'*** is used to close social distance in text type 5, I adopted the negotiation strategy categories from De Dreu, Weingart, and Kwon (2000) who identified two broad approaches to negotiation, i.e., problem-solving (giving information about underlying needs and interests, managing the negotiation process) and contention (positional arguing, substantiation, and making demands). Besides, I also took into account some more detailed strategies used in coding negotiations, such as making demands, priority information exchange, making comments, multi-issue offers (Olekalns & Smith, 2003; Weingart, Thompson, Bazerman, & Carroll, 1990). On these bases, I ultimately decided to put '***you***' from text type 5 into the following categories

1) Asking for something or asking about the others' priorities;
2) Statements of fact which included all the ***you*** pronouns used in factual statements and supportive arguments;
3) Making demands;
4) Showing insights into the others' priorities;
5) Making an offer;
6) other which included all the other ***you*** pronouns used due to grammatical necessity.

The categories of 1), 4) and 5) are conceptually related to closing social distance by focusing on the other party. The other categories of 2) and 3) much less related to closing social distance, while the category 6) could be neutral. After the exploration of the distribution of ***you*** from text type 5 in the five categories, the total occurrence of which is 2548, it is found that about 114 occurrences could be categorized into 1 (asking for something or asking about the others' priorities), 965 occurrences could be categorized into 2 (statements of fact which included all the ***you*** pronouns used in factual statements and supportive arguments), 220 occurrences belong to 3 (making demands), 261 occurrences belong to category 4 (showing insights into the others' priorities), 988 occurrences belong to category 5 (making an offer), and there are not any occurrences categorized into 6. Taken together, those from category 1), 4)

and 5), i.e., totally 1363 occurrences of ***you***, are conceptually related to closing social distance, and those from category 2), 3) and 6) are much less related to closing social distance, which includes 1185 occurrences of ***you***. As illustrated in the following examples, they are from the categories related to closing social distance. Sample 6.57 is an excerpt from category 1, which looks like spoken language, but in a written form; it is more related to a closer social distance. Sample 6.58 and 6.59 are from category 5: sample 6.58 is concerned with making offer, i.e., offering help; in sample 6.59, some advice are given to the readers. In 6.60, the blog writer shows some insights into blog readers' priorities.

6.57) Are **you** a research pro? Do **you** dominate Excel? Are **you** an ardent fan of a current client and already know all about the full product portfolio? (cluster 5, 27.txt)

6.58) If **you** want to venture outside of Milwaukee to visit a farm (and maybe even have breakfast there!), check out this listing of events and dairy farms opening their doors to **you**. (cluster 5, Baderrutter44.txt)

6.59) I suggest **you** shop on the Internet. **You** can get better prices on the internet and send them directly to the bag instead of your home with you to use on your gas at your local store. (cluster 5, Business Property20120413.txt)

6.60) For example, **you** might want to tweet out a link to www.sitbetter.combecause you found a good place to buy office chairs from. (cluster 5, chr20120807.txt)

Apart from the use of ***you*** related to closing social distance, it is also employed in such situations like the statements of fact, making demands, which associate with a relatively greater social distance. Some samples are given below. In 6.61, the blog readers are required to do several things, if they want to explore space. In 6.62, the blog readers are requires to show respect to a person's name. Sample 6.62 and 6.63 are the statements of facts, including some personal opinions as well.

6.61) If you want to make any inroads into exploring space, then **you'**ll surely need to do one or more of three things: become immensely prolific and generative, producing copious amounts of content about yourself; hire my company, Reputation.com, to work with you on maximizing your online reputation profile; or, if you're intent on captivating Google search but really don't have the time, the technical prowess, the writer's gift, or any idea where to start, you can start a reputation club. (cluster 5, chr20120823)

6.62) And you **must** always show respect for the person's time. (cluster 5, LinkedIn20120823.txt)

6.63) The 1-2 days before **you** leave on vacation are very hectic. **You** are doing the last few loads of laundry, picking up, making last minute travel arrangements, running errands, paying bills, and wrapping up any loose ends. (cluster 5, rub1.txt)

6.64) No matter what level in your career, **you** were probably hired because you're a smart individual with applicable skills and knowledge. (cluster 5, Brand Channeler20120627.txt)

Therefore, '***you***' used in cluster 5 could be put into both a social-distance-closing category and a non-social-distance-closing category, although it is more oriented to the social-distance-closing category. It is particularly the dominant use of '***I***', '***we***' and '***you***' that supports Halliday and Martin's (1993, p. 37) argument that ''social context is realized by language''. This will be further discussed in Section 6.4.

6.3.3 Modal verbs

This section presents the results of an investigation of the use of modal verbs across six text types. According to Biber et al. (1999), there are nine central modal auxiliary verbs used to express modality: *can, could, may, might, shall, should, will, would, must*. These modal verbs can be classified into three categories according to their function in language: predictive modals for referring to the future (*will, would, shall*), possibility modals (*can, may, might, could*), which are used to linguistically realize different perspectives; and necessity modals (*ought, should, must*), which directly express persuasiveness (Biber, 1988, p. 241).

A similar procedure taken in the frequency counts of pronouns was followed in this section. The difference is the frequency of each kind of modal verbs was counted, after a list of all the words in the corpus was generated. As can be seen in Table 6.5, the frequency count of each kind of modal verbs in each text type is presented.

Table 6.5: Frequency count of each modal verb in six text types

		Clusters					
		1	2	3	4	5	6
possibility modals	can	146	1192	9	340	751	36
	may	58	277	0	100	108	6
	might	17	74	3	37	68	3
	could	49	172	4	81	85	4
necessity modals	ought	1	0	0	2	0	0
	should	35	214	0	70	107	6
	must	13	116	0	27	24	1
predicative modals	will	171	1006	4	295	490	35
	would	89	353	7	127	106	5

	shall	1	0	0	0	0	0
Total Number of words		103862	284047	3137	91770	96847	11000

When these data were normalized to the occurrences per 10000 Words, the differences among the categories of different modal verbs across six text types were found substantial, as shown in Figure 6.6.

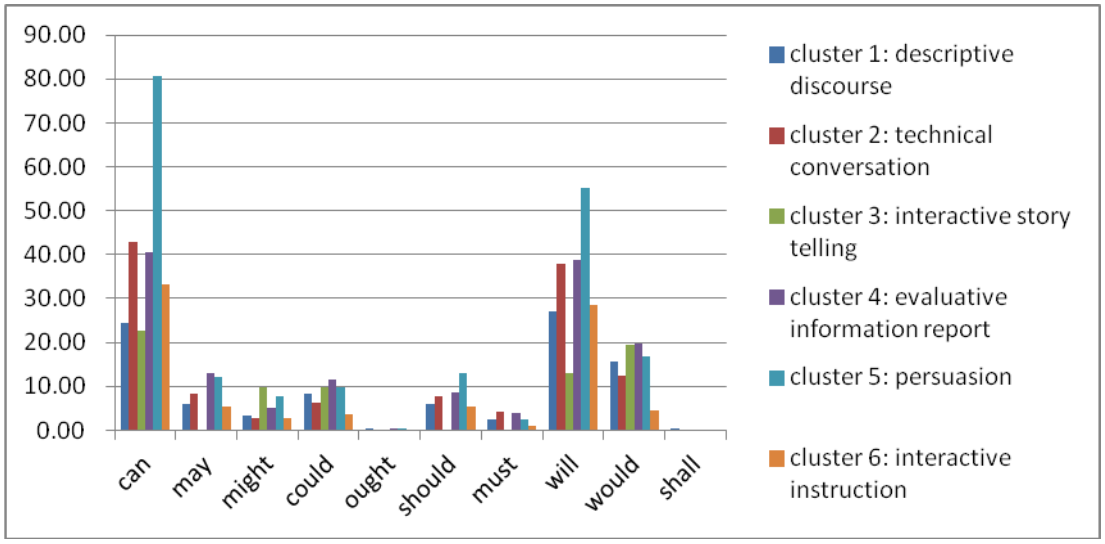


Figure 6.6: Frequency of modals across clusters (Equalized to the Occurrences per 10000 Words)

As illustrated in Figure 6.6, ***can*** and ***will*** are much more frequently used than other modal verbs in most text types apart from text type 3 (interactive storytelling), and both of them are most frequently used in text type 5. In contrast, ***ought*** and ***shall*** are not used at all in several clusters, such as text type 2 (technical conversation), text type 3 (interactive story telling), text type 6 (interactive instruction).

A more macroscopic look at the distribution of modal verbs across different clusters/text types shows that possibility modals are most frequently used, while necessity modals are, over all text types, the least frequent kind of modals, as can be seen in Figure 6.7. This overall picture also shows that there is a highly significant difference between the use of possibility modals and necessity modals in text type 5 (persuasion). It seems that text type5 (persuasion) has a certain notable features undiscovered and it is worthy to be further explored.

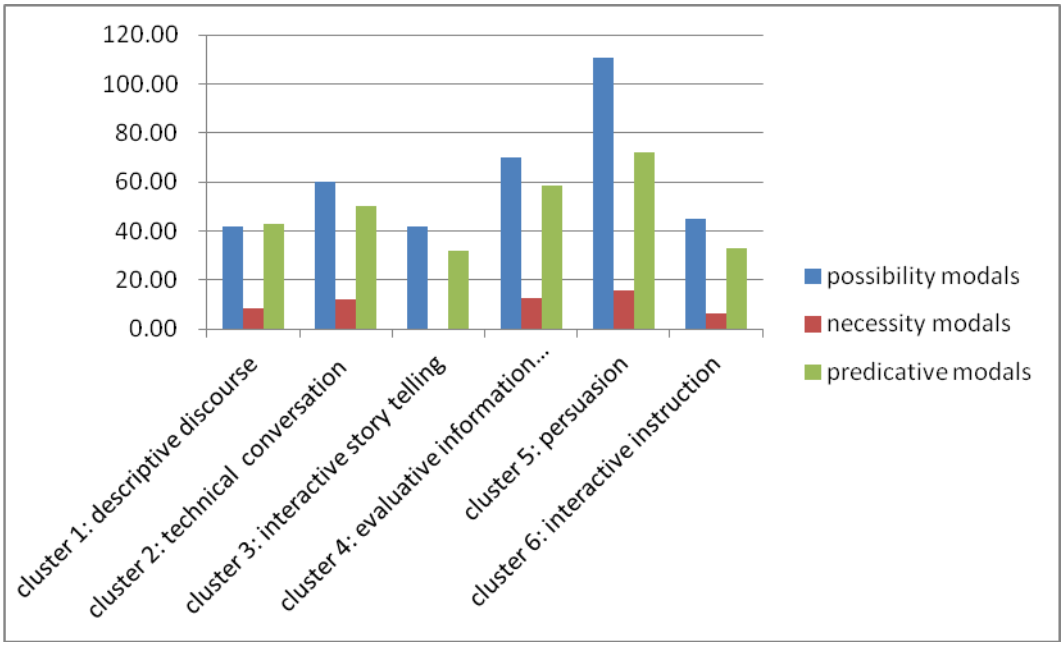


Figure 6.7: Frequency of modal categories across clusters (Equalized to the Occurrences per 10000 Words)

In text type 5 (persuasion), ***can*** is much more dominantly used than other modal verbs. This modal verb can be divided into several categories syntactically as well as semantically. Semantically, like any other modal verbs, ***can*** includes three types: epistemic, deontic, and the dynamic modal of ability (Palmer, 2014). Deontic and epistemic meanings can be labelled as intrinsic modality (meanings relating to permission, obligation, or intention) and extrinsic modality (meanings relating to assessments of likelihood: possibility, necessity) respectively (Biber et al., 1999). Halliday (1970b) subdivides the modals into two basic categories, one of which is the ideational component while the other is part of the interpersonal component. The epistemic modals and some uses of deontic modals (as the use of ***must*** in 6.65) belong to the interpersonal component, because they take care of the speaker's involvement in the utterance: they are "a form of participation by the speaker in the speech event" (Halliday, 1970b, p.335). The modals of dynamic, as well as some other uses of deontic modals (as the use of ***must*** in 6.66), belong to the ideational metafunction, because they "are not speaker's comments, but form part of the content of the clause, expressing conditions on the process referred to" (Halliday, 1970b, p.338).

6.65) You seem to be seeking to destroy yourself in some way, but you ***must*** not include me in your plan of action(Verstraete, 2001).

6.66) But to reach orbit an object ***must*** accelerate to a speed of about 17,500 miles per

hour (Verstraete, 2001).

In 6.65, the deontic modal may also receive reinterpretation just like the epistemic modals, while the deontic modal in 6.66 merely predicates the existence of some kind of necessity without actually committing the speaker to it.

Guided by this framework, it is found that among the 751 occurrences of ***can***, about 300 of them can be categorized into to the ideational metafunction, which indicates the dynamic ability and the epistemic possibility which is usually in the negative form (Coates, 1983). The sample sentences from 6.67 and 6.68 illustrate this. In 6.67, ***can*** is used to indicate people's real talent and ability. In 6.68, ***can*** is an epistemic modal meanings dealing with the possibility of an inference drawn from available evidence.

6.67) The actual bit shuffling in cryptography is the "easy" part of security. We **can** prove lots of things about entropy and ultimate security of the data. (cluster 5, BreakingPoint20120605.txt)

6.68) Events have become a staple in the PR arsenal, and when used effectively, they **can** be an extremely valuable tool for driving awareness and visibility for a brand. (cluster 5, Brand Channeler20111025.txt)

In contrast, a relatively greater number of ***can*** frequently collocates with subject ***you,*** representing a personal involvement and interpersonal function in the clause, which encodes what possibly will, or should be implemented, as can be seen in sample sentences from 6.69 to 6.70. All these examples focus on making offers, and suggestions; they are the variation of 'command' ***may***, exhibiting an extended meaning of the permission.

6.69) Once you have your Economy ticket purchased, you **can** choose an Economy Comfort seat via delta.com, our self-service kiosks or Delta reservations for an additional fee of $19-$99 one-way including Alaska, Hawaii, and short-haul international flights. (cluster 5, Delta20111020.txt)

6.70) With Cloud Search, you **can** focus on your application layer. You upload your documents, Cloud Search indexes them, and you **can** build a search experience that is custom-tailored to the needs of your customers. (cluster 5, Amazom20120412.txt)

In comparison with the predominance of ***can,*** which is a kind of possibility modal, necessity modals, i.e., ***ought***, ***should***, ***must***, are much less frequently used. ***Must*** is the kind of modal verb indicating the strongest obligation and necessity. There are a wide range of subjective as well as objective uses of modal expressions across epistemic ***must*** and deontic ***must***. The subjective use of epistemic ***must*** may range from being based one the speaker's personal

knowledge to being based on common sense, which conveys the speaker's confidence in the truth and the facts known to him; the subjective use of deontic ***must*** could be fulfilled by the following conditions: the subject is animate, the main verb is an activity verb, the speaker is interested in performing the action, the speaker has authority over the subject (Coates, 1983). In contrast, the objective use is extremely uncommonly used in the everyday communication (Lyons, 1977).

Among the total 181 occurrences of ***must***, 116 of them are from cluster 2 (technical conversation), and 24 of them are from cluster 5 (persuasion). The total occurrences of ***must*** were classified according to the animate-inanimate agent division. It was found that 32 out of 116 occurrences of ***must*** from text type 2 (technical conversation), and 8 out of 23 occurrences of ***must*** from text type 5 (persuasion) have an animate agent. This result shows that the use of the modal verb ***must*** in both text type 2 (technical conversation), and text type 5 (persuasion) is to share some personal knowledge or common sense (e.g., 6.71), though occasionally giving an authorative command (e.g., 6.72), as shown in the following examples:

6.71) Remember that tickets need to be reissued on or before and travel **must** occur no later than January 13. (cluster 5, Delta20110107.txt)

6.72) However, you **must** make sure you test for the exact conditions highlighted above and consider what is important to a successful deployment of CGN. (cluster 2, BreakingPoint20120214.txt)

In tracing the use of ***must*** and their relationship to the relative power between the writer and the reader, i.e. the comparison between the blogs categorized into different text types, there does not seem to be a great deal of difference in the frequency of ***must*** in the two text types. Overall, regarding the role that social distance plays in the directness of expressing obligation, there is not too much evidence found in the case of ***must*** in terms of their distinction across different text types, apart from the finding that necessity modals, as a whole, are much less used in corporate blogs***.*** In order to further investigate how the CB text types distinguish with each other and the overall characteristics of corporate blogs, both lexical density and grammatical complexity are explored, and the results are presented below.

6.3.4 Lexical density

In this section, the use of lexical density (LD) across six text types was compared. The measurement of LD in each text type followed Halliday's formula (Halliday, 1985), by means of which the proportion of the number of content words to the total running word is calculated. A difficult thing in the calculation is the categorization of leixcal items (content words), and grammatical items (function words), both of which constitute a complete lexicogrammar system. Although lexical items and grammatical items have been determined differently by linguists and researchers in literature (cf. Chapter 2), I adopted Halliday's (1996, p. 347) measurement to guide this analysis, which calculates "the proportion of content (lexical) words - basically nouns, full verbs, adjectives, and adverbs derived from adjectives - over the total number of words in a text". The data prepared for this calculation are six directory files, representing six text types of corporate blogs, in which tagged texts rather than raw texts are contained. A similar procedure taken in the frequency counts of pronouns and modal verbs was followed in this section. The difference in the calculation of lexical items (content words) in this analysis is more complex. Considering that content (lexical) words include nouns, full verbs, adjectives, and adverbs derived from adjectives, the following tagsets were included, as can be seen in Table 6.6. Among the tags associated with verbs, those representing auxiliary verbs and linking verbs are not included. The auxiliary (helping) verb is a type of non-lexical verb, with mostly grammatical functions helping to form complex verbal groups. Specifically, the modal verbs, e.g., *will, would, shall, should, can, could, may, might*, and some auxiliary verbs, e.g, *do*, *have*, which function as predicate together with the past participle or the infinitive of a full verb are not included. Different forms of the linking verb *be* are not included either. Besides, some tags which may either refer to the main verb (full verb), or the auxiliary (helping) verb, i.e., different forms of *do,* and *have*, are included. Only those functioning as the main verbs should be retained; therefore, their identification depends on their use in the context. Overall, the verbs that cannot be categorized into full verbs are not included in the calculation of content words. Presented below is a table which shows the tags included in the calculation of content (lexical) words.

Table 6.6: Tags included in the calculation (from CLAWS7 (C7) POS Tageset

http://ucrel.lancs.ac.uk/claws7tags.html)

Part of speech	Tags	Examples
adjectives	JJ	general adjective
	JJR	general comparative adjective (e.g. older, better, stronger)
	JJT	general superlative adjective (e.g. oldest, best, strongest)
	JK	catenative adjective (able in be able to, willing in be willing to)
nouns	ND1	singular noun of direction (e.g. north, southeast)
	NN	common noun, neutral for number (e.g. sheep, cod, headquarters)
	NNA	following noun of title (e.g. M.A.)
	NNB	preceding noun of title (e.g. Mr., Prof.)
	NNL	singular locative noun (e.g. Island, Street)
	NNO	numeral noun (e.g. dozen, hundred)
	NNT	temporal noun (e.g. day, week, year)
	NNU	unit of measurement (e.g. inch, feet)
	NP	proper noun (e.g. IBM, Andes)
	NPD	weekday noun (e.g. Sunday)
	NPM	month noun (e.g. October)
Verbs	VD0	do, base form (finite)
	VDD	did
	VDG	doing
	VDI	do, infinitive (I may do... To do...)
	VDN	done
	VDZ	does
	VH0	have, base form (finite)
	VHD	had (past tense)
	VHG	having
	VHI	have, infinitive
	VHN	had (past participle)

	VHZ	has
	VV0	base form of lexical verb (e.g. give, work)
	VVD	past tense of lexical verb (e.g. gave, worked)
	VVG	-ing participle of lexical verb (e.g. giving, working)
	VVGK	-ing participle catenative (e.g. going in be going to)
	VVI	infinitive (e.g. to give... It will work...)
	VVN	past participle of lexical verb (e.g. given, worked)
	VVNK	past participle catenative (e.g. bound in be bound to)
	VVZ	-s form of lexical verb (e.g. gives, works)
Adverbs	R--	words tagged with R initiated tags, adverbs derived from adjectives

And the frequency count of each lexical item in each text type is presented in Table 6.7, as well as Figure 6.8.

Table 6.7: Lexical density across six clusters/text types

	Cluster 1: descriptive discourse	Cluster 2: technical conversation	Cluster 3: interactive story telling	Cluster 4: evaluative information report	Cluster 5: persuasion	Cluster 6: interactive instruction
A= Number of lexical words	24562	136554	1473	40649	41241	5350
B=Total number of words (tokens)	103719	284047	3137	91770	96847	11000
LD (%) = A/B x100	23.68	48.07	46.96	44.29	42.58	48.64

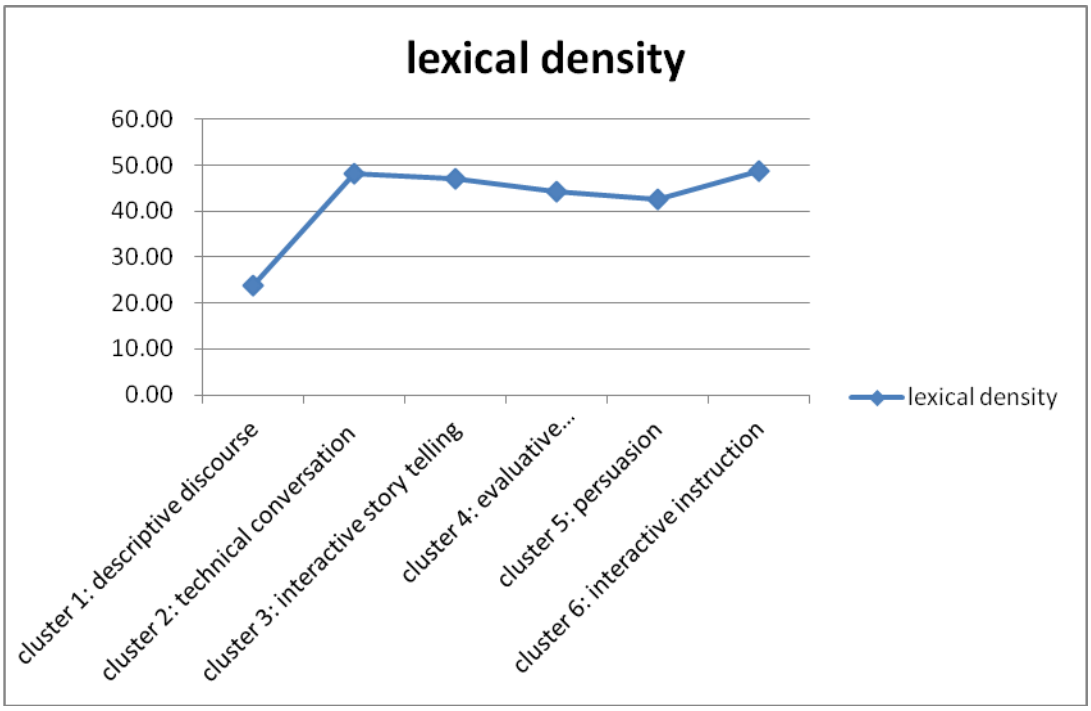


Figure 6.8: lexical density across six clusters/text types

In Figure 6.8, the difference between the maximal and minimal degrees of lexical density was not very broad: the lexical density is 23.68% for text type 1 (descriptive discourse), and 48.64% for text type 6 (interactive instruction). They are the lowest and the highest one respectively. The LD in the other text types is relatively close to each other. Based on Ure's (1971) claims that the spoken texts have a lexical density of less than 40% (ranging from 24% to 43%) and the written texts a density of greater than 40% (ranging from 36% to 57%), a conclusion can be drawn that the corporate blog is primarily in the written form, consisting of both spoken and written modes, since the LD in most CB text types are greater than 40%.

In a more general sense, Halliday and Martin (1993) contend that lexical density is indicative of information density in a text, which could be measured in accordance with "how tightly the lexical items have been packed into the grammatical structure" (Halliday & Martin, 1993, p. 76). It can be measured by the number of lexical words per clause. By using two measures from the fourteen syntactic/grammatical complexity measurement (Lu, 2010) (cf. Section 6.3.5), the ratio of the number of lexical words to clauses was calculated. The formula is presented below:

Step 1 = the ratio of the number of lexical words to the total number of words (cf. Table 6.7)

Step 2 = the ratio of the number of lexical words to the total number of words ×MLC (# of words / # of clauses) (cf. Table 6.8)

By using this formula, I conducted another LD calculation (LD2), and got the following results: the ratio of the number of lexical words to the clauses are 2.91 (approximately 3 lexical items and 1 ranking clause), 5.77 (approximately 6 lexical items and 1 ranking clause), 4.18 (approximately 4 lexical items and 1 ranking clause), 4.89 (approximately 5 lexical items and 1 ranking clause), 4.42 (approximately 4 lexical items and 1 ranking clause), and 6.13 (approximately 6 lexical items and 1 ranking clause) respectively across six text types. Putting the results of the two methods of the calculation of lexical density in a graph (Figure 6.9), it can be seen that, by using two approaches, the results of lexical density are quite similar. Text type 1 (descriptive discourse) has the lowest lexical density than other text types because there are more ranking clauses to share the lexical items, while text type 6 (interactive instruction) has the highest lexical density. As claimed by Halliday (1985) , 'typically, written language becomes complex by being lexically dense: it packs a large number of lexical items into each clause'. This result basically corresponds with the previous LD analysis.

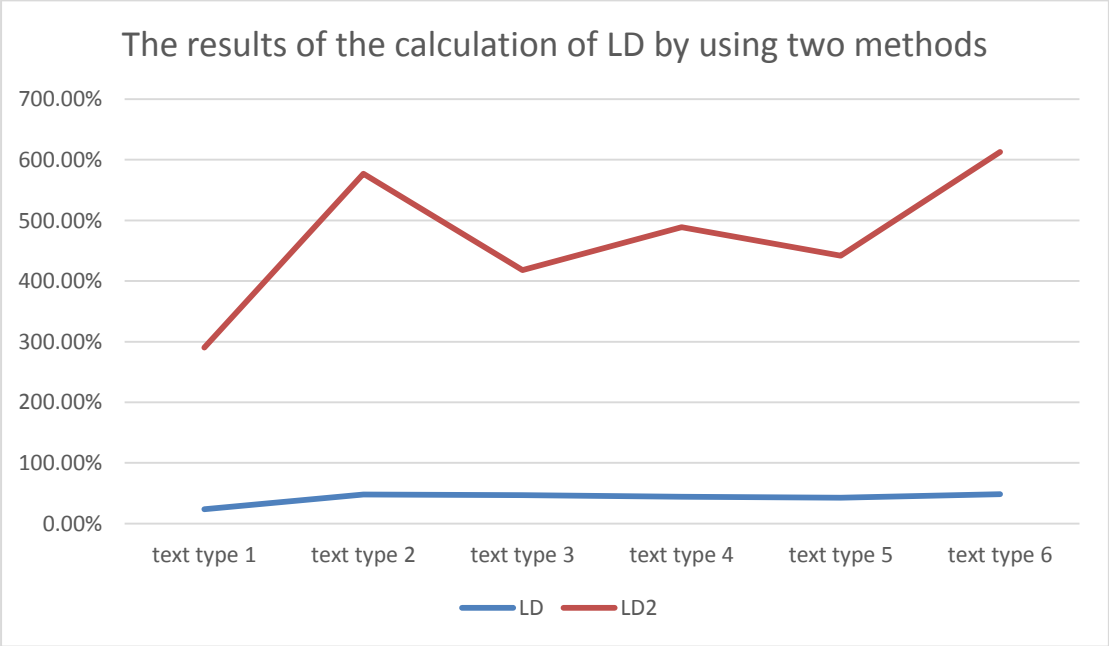


Figure 6.9: The results of the calculation of LD by using two methods

As for the explanation for greater lexical density, Halliday and Matthiessen (2004, p. 655) have claimed that "the nominal group is the primary resource used by the grammar for packing in lexical items at high density". As claimed by Leech (2009), the greater lexical density results from the compression of information for some reasons, e.g., the pressure to condense as much information as possible into the space available in the press. If taking text type 6 as an

example, it could be found that there is a frequent occurrence of the noun (N) + noun (N) combination, as shown in the sample excerpts in Figure 6.10 and Figure 6.11:

Figure 6.10: N+N combinations in text type 6-1

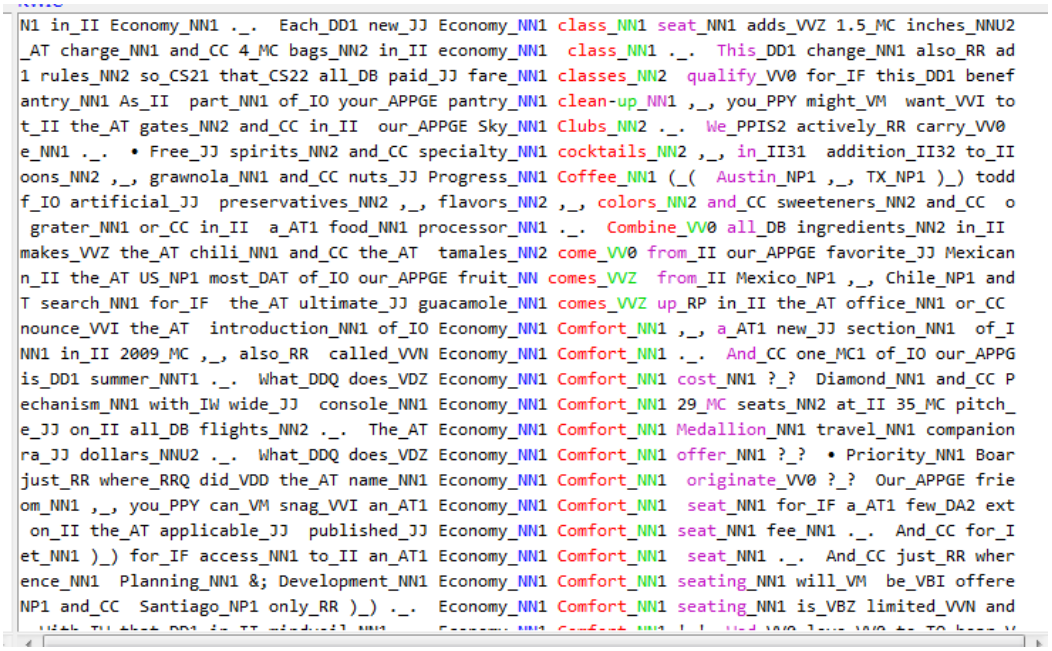


Figure 6.11: N+N combinations in text type 6-2

In comparison with text type 6 (interactive instruction), text type 1 (descriptive discourse) involves a relatively smaller number of N+N combinations, as shown in Figure 6.12. Statistically, among the total number of nouns tagged as NN in text type 1 (descriptive discourse), i.e., 15262 nouns, 1637 of them co-occur with another noun, constituting a 10.7% of N+N combination. In text type 6 (interactive instruction), this number is 21.3% (603/2825).

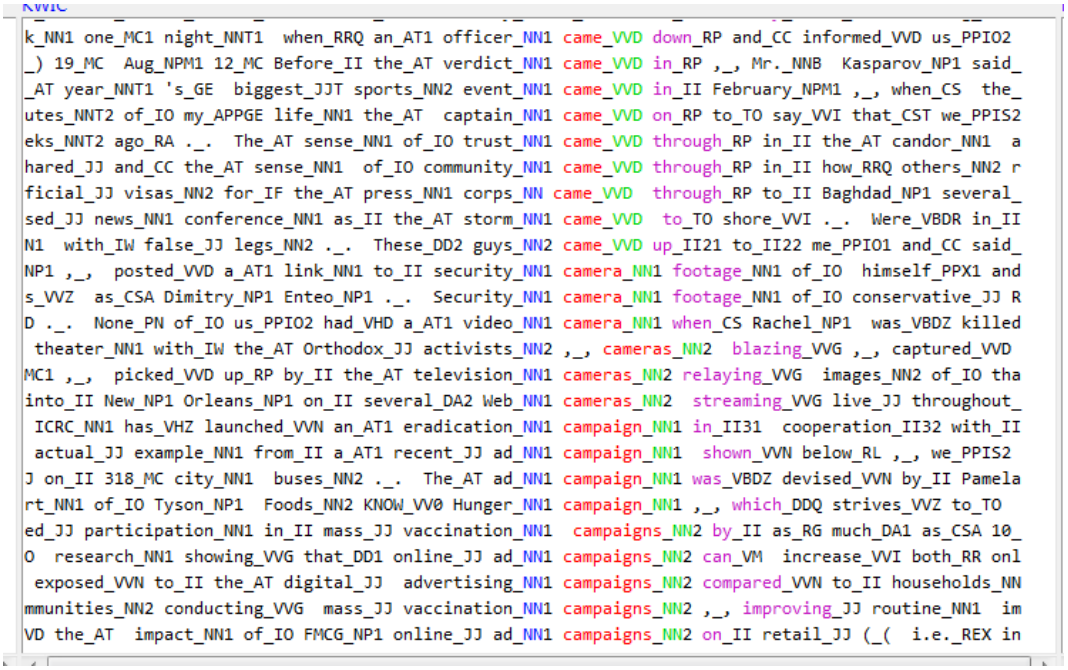


Figure 6.12: N+N combinations in text type 1

This comparison shows that there are more condensed information in text type 6 (interactive instruction) than in text type 1 (descriptive discourse), which corresponds with the previous finding that text type 1 (descriptive discourse) is closer to spoken texts, while text type 6 (interactive instruction) is closer to written texts. However, all the text types in corporate blogs are in the written form, constituting a dynamic as well as a constant reservoir. Overall, LD is probably more indicative of a colloquial-literary continuum rather than of the difference between spoken/written medium. It can also be noted that the style and genre may have a greater effect on textual differences, in particular, LD increasing and LD lowering features should be linked to literary/written and colloquial/oral styles respectively rather than the oral/written medium of the text.

6.3.5 Grammatical complexity

This section presents the result of the exploration of grammatical complexity across six text types. By implementing a syntactic complexity analyser[16] designed by Lu (2010) which is manipulated in python and runs on UNIX-like systems with Java 1.6 and python 2.5 or higher installed, the 14 grammatical complexity features are computed, which include Mean length of clause (MLC), Mean length of sentence (MLS), Mean length of T-unit (MLT), Sentence

[16] The analyzer is called the Stanford parser (Klein & Manning, 2002) for parsing the input file and Tregex (Levy & Andrew, 2006) to query the parse trees

complexity ratio (C/S), T-unit complexity ratio (C/T), Complex T-unit ratio (CT/T), Dependent clause ratio (DT/C), Dependent clauses per T-unit (DC/T), Coordinate phrases per clause (CP/C), Coordinate phrases per T-unit (CP/T), Sentence coordination ratio (T/S), Complex nominals per clause (CN/C), Complex nominals per T-unit (CN/T), Verb phrases per T-unit (VP/T) (details see Table 4.4). As can be seen in Table 6.8, Table 6.9, as well as Figure 6.13, and Figure 6.14, the mean scores of the 14 grammatical complexity features in the 6 text types are presented.

Table 6.8: Grammatical complexity values across six text types-1

Grammatical complexity features (mean values)	Cluster 1 descriptive discourse	Cluster 2 technical conversation	Cluster 3 interactive story telling	Cluster 4 evaluative information report	Cluster 5 persuasion	Cluster 6 interactive instruction
MLS	25.64	22.52	17.89	22.5	20.72	20.26
MLT	24.83	20.57	16.23	20.6	18.54	19.31
MLC	12.27	12.00	8.91	11.03	10.37	12.6
C/S	2.21	1.95	2.02	2.08	2.08	1.64
VP/T	2.51	2.35	2.31	2.48	2.51	2.12
C/T	2.13	1.76	1.83	1.89	1.85	1.55
DC/C	0.33	0.34	0.36	0.39	0.39	0.32

Table 6.9: Grammatical complexity values across six text types-2

Grammatical complexity features (mean values)	Cluster 1 descriptive discourse	Cluster 2 technical conversation	Cluster 3 interactive story telling	Cluster 4 evaluative information report	Cluster 5 persuasion	Cluster 6 interactive instruction
DC/T	0.68	0.63	0.67	0.77	0.73	0.52
T/S	1.07	1.10	1.1	1.1	1.12	1.05
CT/T	0.42	0.44	0.47	0.52	0.49	0.38
CP/T	0.45	0.56	0.37	0.53	0.47	0.63
CP/C	0.28	0.34	0.2	0.29	0.28	0.43
CN/T	2.62	2.29	1.53	2.43	1.84	2.23
CN/C	1.37	1.34	0.83	1.31	1.04	1.45

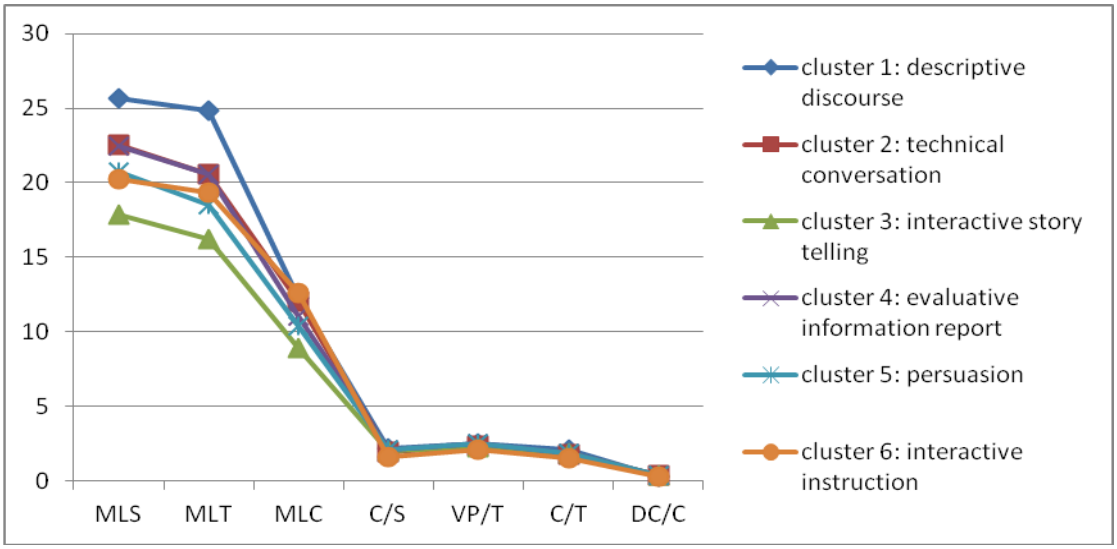


Figure 6.13: Mean scores for grammatical complexity analysis across six text types-1

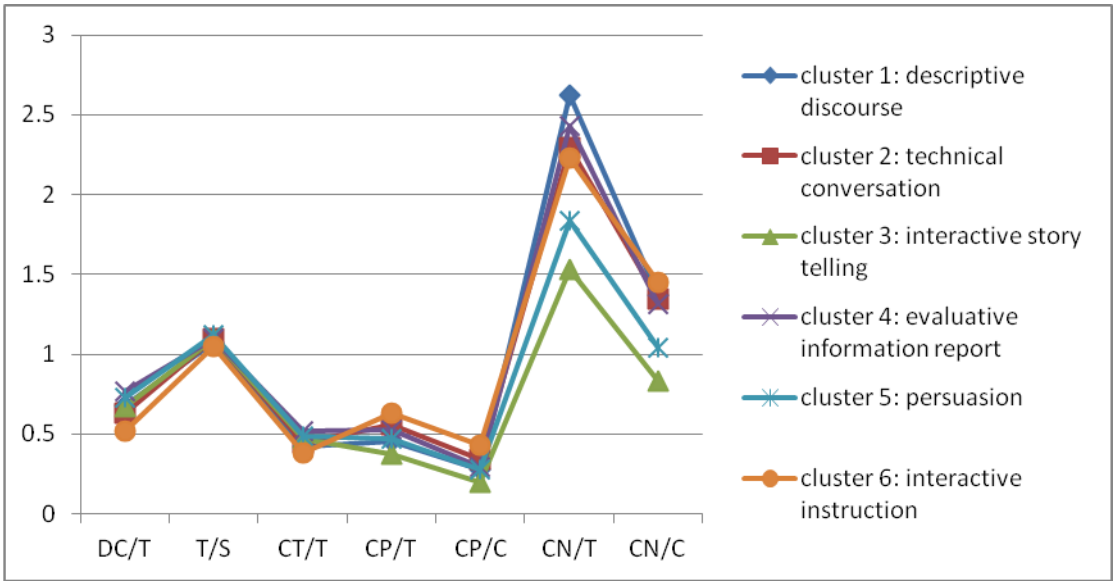


Figure 6.14: Mean scores for grammatical complexity analysis across six text types-2

In Figure 6.13 and Figure 6.14, it is shown that most measures are very close to each other, apart from MLS (Mean length of sentence), MLT (Mean length of T-unit), CN/C (Complex nominals per clause) and CN/T (Complex nominals per T-unit). Among the four grammatical complexity features that are divergent to each other across different text types, it is in text type 1 (descriptive discourse) that involves the longest MLS, MLT, and the most complex CN/C and CN/T, and text type 3 (interactive story telling) that involves the shortest MLS, MLT, and the least complex CN/C and CN/T. Each CB text has a grammatical complexity value in terms of MLS, MLT, CN/C and CN/T, which vary dramatically from each other, and the MLS, MLT, CN/C and CN/T identified in both text type 1 (descriptive discourse) and text type 3 (interactive story telling), as well as in other text types are the mean value. Presented below are some text

excerpts from the texts with a value of MLS somewhat higher or lower than the mean value from text type 1 (descriptive discourse), text type 3 (interactive story telling) and text type 5 (persuasion). It can be seen that the MLS in 6.73 is somewhat higher than that in 6.74, and much higher than that in 6.75. Obviously, the CB text 6.73 which has the highest MLS is the most complex grammatically, while 6.75 has the least grammatical complexity. This is the same case to MLT, CN/C and CN/T, i.e., the higher the value of MLT, CN/C or CN/T is, the more complex a CB text or a text type is grammatically, since it has been assumed and proved that T-units contain more nominal structures in written than spoken language, and there is a greater average length of T-units in written language (Biber et al., 2011; Hudson-Ettle, 1998).

6.73) For the Educators' Challenge, teachers nationwide were tasked with creating a detailed action plan on how they would promote safe driving among teens in their school and in their community. This year's winners were chosen for their plan's creativity, content and structure. First Place winner Erin Hattabaugh from Cleveland High School in Cleveland, Tenn., has won a Toyota Driving Expectations in-school safe driving event, a virtual driving simulator for her school and a $5,000 grant to implement her winning action plan. Second Place winner Carrie Shampoe from Clymer High School in Clymer, N.Y., has won a virtual driving simulator for her school and a $1,500 grant to implement her winning action plan. Both winning lessons will be made available on the Toyota Teen Driver website for teachers to use. (cluster 1, toyota20120504.txt, MLS=117.8333)

6.74) Frozen and canned vegetables are a great option. We wouldn't have made it, frankly, without them. We found mixed veggie combos to be delicious, taste fresh, and used them liberally in stir-fries, pasta sauces and just about everywhere.

The more you cook from scratch, the cheaper it is. I made two loaves of bread every week for a mere $1.25. It makes sense: if you do the labor, you keep the savings. This is pretty true for just about everything in the store. (cluster 3, WholeFoods20120222-2.txt, MLS=16.5227)

6.75) The following instruction is for Eclipse Indigo.

1. Download Java ME SDK and Eclipse plugins from here.
2. Install Java ME SDK
3. Launch Eclipse
4. Go to "Help" - "Install New Software"
5. Click "Archive", and choose the Eclipse Plugin zip file you downloaded. Name is something like "Java ME SDK Plugin" (cluster 5, ora20120925-21.txt, MLS=8.175)

Taken together, as a mean value, MLS, MLT, CN/C and CN/T are the highest in text type 1 (descriptive discourse) and the lowest in text type 3 (interactive story telling), representing a conclusion that text type 1 (descriptive discourse) and text type 3 (interactive story telling) are the most peculiar text types in corporate blogs: text type 1 (descriptive discourse) is the most

complex grammatically, while text type 3 (interactive story telling) has the least grammatical complexity. This complexity constitutes a text pool, in which a variety of text types are included and different communicative purposes can be achieved. The CB texts might be a personal story, a persuasive text, or an introduction to a particular product, the purposes of which could be customer relationship management, or for sale. When it comes to different kinds of participants involved, the CB texts could be prepared for the existing or potential customers, to which relevant and valuable information are provided.

6.4 Discussion

In this chapter, some lexical and grammatical features (i.e., the most common verbs, pronouns, modal verbs, lexical density and grammatical complexity) associated with Halliday's metafunctions have been explored, and compared across six text types. These analyses have revealed not only quantitative but qualitative differences in the use of these language features.

Quantitatively, there is a significant difference in the use of these language features across six text types. The most notable results are: 1) ***said, using, found, using, need, want*** are respectively the most common verbs in each of the six text types; among the fifteen most common verbs in CBC, the verb ***make*** is the only one that takes all the tenses (***make, makes, made, has/have made, making***) and patterns, thus constituting a great proportion of occurrences. 2) There is a dominant use of '***I***' and '***we***' in text type 3 (interactive storytelling), '***you***' in text type 5 (persuasion), '***can***' and '***will***' in most text types excluding text type 3 (interactive storytelling), and a rare use of necessity modal verbs in all CB text types. 3) most CB text types are close to each other in terms of their lexical density and grammatical complexity, apart from text type 1 (descriptive discourse). Text type 1 has the lowest LD, the longest MLS, MLT, and the most complex CN/C and CN/T, which means text type 1 (descriptive discourse) is grammatically complex but lexically simple. Taken together, text type 1 (descriptive discourse) could be seen as the most spoken text type in corporate blogs, since spoken language is often complex grammatically and simple lexically (Halliday, 1989).

In terms of the exploration of the most common verb and process types in corporate blogs, it

illuminates the way these frequently used verbs construe the world of experience into a manageable set of process types in this genre. In particular, in each of the six text types, there is one highest frequency verb identified. The highest frequency of ***said*** (verbal) in text type 1 (descriptive discourse) shows that the descriptive discourse is most represented by providing the reader what others have said or has experienced, thus making the reader as vividly aware as possible of what the writer has perceived and experienced personally through his senses or activities. Likewise, both of text type 2 (technical conversation) and text type 4 (evaluative information report) exhibit the highest percentage for the use of ***using*** (material). A plausible reason for this might be that the communication process, in which the creation and the delivering of a safe, efficient and effective use of products or services, is focused on. The content delivered in a blog may range from context-sensitive topics to user-personalized help. And it is the frequent use of ***using*** that suggests a strategy employed by the blog writers, i.e., considering that consumers would rather scan for the relevant information they need at hand than read the blog posts from start to the end, blog writers then present some potentially important or urgently needed content by focusing on some specific issues, such as the use of products, which is supported by a frequent use of ***using***. Besides, the use of ***need*** in text type 5 (persuasion) and ***want*** in text type 6 (interactive instruction) has some overlap with each other. When the verb ***need*** in text type 5 (persuasion) is characterized by the consciousness of the Sensor who wants something and thus categorized into mental process (31 out of 114 ***need***), it is very close to the use of ***want*** in text type 6 (interactive instruction), both of which are concerned about some requirements and intentions. However, the difference is text type 5 (persuasion) is more interactive than text type 6 (interactive instruction), in which the verb ***need*** mainly co-occurs with ***you***. That means when a message is sent from one party to another with a focus on ***you***, it attempts to reinforce their attitudes, and influence the target audiences, though the outcome may not be guaranteed. In contrast, the verb ***want*** in text type 6 (interactive instruction) only co-occurs with the first person pronouns, including ***I*** and ***we***, which suggests a main focus on the message from a blog writer's perspective; accordingly, the readers are less influenced by some extra factors other than the contents of the message. At last, text type 3 (interactive story telling) consists of the smallest number of CB texts, and the verb ***found*** is used vividly and precisely to get the reader involved in the elements of the story,

which includes plot, character, setting, etc.

Apart from the identification of the most common verbs in each of the six text types, their process types were identified as well. Although the current analysis does not provide a comprehensive picture of the categorization of process types in corporate blogs, the preponderance of the top fifteen verbs directs our attention to the most prototypical activity in corporate blogs. In terms of the process types of the most frequent fifteen verbs, they have 3246 occurrences and could be categorized into three process types. Among the top fifteen most common verbs, there is a preponderance of material process (1961 out of 3246 material process verbs, accounting for 60.40%), followed by mental process (998 out of 3246 mental process verbs, accounting for 30.75%) and then the verbal process (287 out of 3246 verbal process verbs, accounting for 8.84%). This result shows that by using material processes, a writer can represent both a subjective and an objective experience related to an instruction/procedure. With this process constituting a high percentage among all the processes in the most common verbs, Halliday and Matthiessen's (2004) viewpoint that material processes are one of the three basic types seems pertinent. On the other hand, a high verbal process frequency which was found particularly in text type 1 (descriptive discourse) is partly because this text type is associated with the use of a conversational and casual style, but also due to the authors' writing experience. Besides, with desiderative mental process-verbs 'want' and 'need' most frequently used in text type 5 (persuasion) and text type 6 (interactive instruction), it significantly expresses the inner desire or the want of either the blog writer or the blog readers. That is, these desiderative mental processes have the function of divulging the obsession each person is indulged in, and similarly reveals each person's intention.

Also presented below is a comparison between the most common verbs identified in corporate blogs and in Biber et al.'s (1999) comprehensive study on LSWE corpus. In Biber et al. (1999), it is found that there are 12 very common verbs unevenly distributing across different semantic domains, among which there are six activity verbs (***get, go, make, come, take, give***), five mental verbs (***know, think, see, want, mean***), one communication verb (***say***), the single most

common lexical verb overall. Most of these words are also frequently used in the CBC, which include ***said*** (past tense of ***say***), ***know,*** different forms of ***make*** (***making, make, makes, made***), ***want***, and ***seen*** (past perfect tense of ***see***). A notable difference between the CBC and LSWE corpus is they have their different most common verb respectively: the verb ***say***, which is commonly used in conversation, is the most frequent verb overall in the LSWE Corpus and the only verb to be extremely common in more than one register. In the CBC, however, ***say*** is much less used than its past tense form (***said***); rather, ***using*** is most frequently used in corporate blogs, particularly in the text types 2 (technical conversation) and text type 4 (evaluative information report). These commonly used material process verbs in CBC tend to be pragmatically associated with each other. Most of them reflect commonly shared topics, which are to a very large extent concerned with processes essential for reporting the physical activity in a business context. That is, the material processes could be positioned as signifying partnership qualities of sharing activity information (e.g. to make, to work, to use), and narration concerning customer care (e.g. to provide, to help).

Related with interpersonal metafunction, it is found that '***I***', '***we***' and '***you***' are more frequent used than other personal pronouns. This is in conformity with the personal principles claimed in Maslansky et al. (2010). Besides, it does not totally coincide with Maslansky et al. (2010) in the way that it does not show a marked preference on ***I*** or ***we***; rather, the total occurrences of ***you*** is much more than ***I*** or ***we*** (cf. Table 6.4). No matter what kind of person pronouns are preferred, the investigation of them in CBC shows the author's strategic decisions in communication which affect distance between authors and readers. Since the strategic decisions in different text types are varied, there is a tension created and adjusted dynamically to balance the social distance between authors and readers, thus achieving the communicative purposes of these text types. In particular, '***I***', '***we***' and '***you***' are all associated with the blog writer's attitude towards the readers. The vast majority of '***we***' employed in corporate blogs is exclusive-***we***, without including the target audiences/readers; however, this does not mean the writers intend to show a social stance, or avoid an in-group solidarity with the readers. Rather, the dominant use of exclusive-***we*** in the single-authored CB texts is, to a large extent, context based, i.e., it indicates an intention to speak as an authority and reduce

personal attribution. In text type 6 (interactive instruction), the use of exclusive-***we*** shows a large power relation, since the primary goal of giving an instruction is presenting information to someone for guiding them or teaching them to do something, which usually involves the steps that must be followed. However, the use of interactive language in this text type may help reduce some knowledge-based authorities of blog writers and even simultaneously establish close rapport with blog readers. In this way, corporate blogs can be put further to a spoken end in a spoken-written continuum. In terms of ***you***, and there is not a clear-cut distinction among them apart from the use of '***you***' in text type 5 (persuasion). '***You***' could be put into both a social-distance-closing category and non-social-distance-closing category, mostly representing the writer him/herself as distanced from the readers and from the phenomena as a neutral observer based on his/her knowledge superiority. There is also an uncertainty that the second personal pronoun '***you***' may refer to a single person or the general public, or a certain group of people. It is obvious that the frequent use of '***you***' in text type 5 (persuasion) shows that the blog writer aims to send messages, with the hope of bonding with his audiences, as well as gaining trust from them. Besides, the big contrastive use of possibility modals and necessity modals in text type 5 (persuasion) shows that blog authors have been trying to avoid the use of the language of command, or creating obligations when they try to persuade their readers to accept a new product, or probably a new policy. Rather, they hope to convince their audiences by means of making suggestion, and presenting logical possibility or ability to be acquired. Similarly, even in those more complex CB texts, more complex structures, rather than large amount of technical words or terminologies were used, the purpose of which is to attract people with or without any expertise.

In terms of the mode of discourse, the degree of lexical density and grammatical complexity is highly sensitive to the variation of mode, and they are also responsible for the differences between spoken and written languages. The exploration of these two language features shows that the highest level of lexical density and grammatical complexity occurs in the text type of description discourse, followed by other text types which do not distinguish dramatically from each other in terms of their LD and GC. Overall, the corporate blog is primarily a written discourse, in which text type 1 (descriptive discourse) is the most involvement-oriented spoken

text type. Within each text type, there are possibly some strategies employed by the companies to balance the social power from the expertise to the individuals, and the maintenance of a warm and friendly online atmosphere. In this way, the main organizational goals of corporate blogs are achieved; at the same time, a good buyer-supplier/ writer-reader relationship is established. In CB texts from text type 2 (technical conversation), the technical information is delivered to their audiences in a conversational tone; in CB texts from text type 5 (persuasion) as well as the e-commerce and online social networking service industry, the audiences are persuaded to do something. The blog writers tell some personal stories to their audiences in text type 3 (interactive storytelling), or report some information very objectively in the CB texts from text type 1 (descriptive discourse) and the publishing industry; they provide some instructions to their audiences in the CB texts from text type 6 (interactive instruction) and the whole sale and retail industry. In particular, storytelling rooted in the reality of products and industry, or blog writers' personal experiences could be considered as a strategy to build image and branding, as it is able to increase consumer emotional involvement in the brand. Some other text types tend to share information with the readers, but show different social powers and distance. CB texts from text type 2 (technical conversation) and text type 6 (interactive instruction) show a high social power characterized by people's perception in knowledge and trustworthiness. When blog writers, on behalf of their companies, share some knowledge or instructions as a credible one, others would be convinced. In this way, new sales or services could be conducted directly to the audiences/readers. In comparison with this, text type 5 (persuasion) is much more ambitious, which aims to persuade the audiences to accept their ideas or agree to a specific activity, such as buying their services or products. In contrast, text type 1 (descriptive discourse) tends to simply explain, describe, or inform. Taken together, corporate blogs serve various purposes, which include building image and establishing community between readers and writers, developing new product, conducting direct sale to a wide range of audiences, and helping solve problems by giving instructions and explanation, etc.

Comparing with the previous studies on corporate blogs (Puschmann, 2010a) and on other computer-mediated discourse from a linguistic perspective (Grieve et al., 2011; Herring, 2001;

Herring, 1996; Stuart, 2006; Yates & Orlikowski, 1993; Yates, 1996), it can be seen that the current research provides a plausible and extra perspective to look at corporate blogs as a register. Previous research has, in general, agreed upon the 'hybrid' nature of CMC with both oral and written discourse features, such as the study on instant messaging (Ferrara, Brunner, & Whittemore, 1991), Internet Relay Chat (Werry, 1996), etc. It is also claimed that drawing a clear line between the spoken and the written mode in any CMC types, such as email (Baron, 1998), is a complex issue. Besides, it is also acknowledged that the language in CMC varies according to functions in contexts, and vice versa, extra-linguistic factors influence the language of CMC (Baron, 1998; Hård af Segerstad, 2002). The current approach to examining corporate blogs differs from previous research through its focus on the register variation reflected by the co-occurring linguistic features, as well as the salient linguistic differences across them in the configuration of the three components of the context of situation - field, tenor, and mode. As such, the analysis offers a fuller and more comprehensive examination of the ways in which the register variation in corporate blogs functions in the field of discourse, the mode of discourse and the tenor of discourse. As for this study, some more general conclusions can be made.

Firstly, the construction of field, tenor, and mode was complex and interrelated, which involves not only a linguistic, but also a social and cultural process. For blog writers, there is always a dilemma between the intentions of promoting their company and products, and creating a supportive, respectful atmosphere. In contrast to using blogs as a direct sales tool, those who share their personal stories in blogs also face challenges, since it is not sure whether or not their stories are anticipated by the readers, unless they are big names. Therefore, these top-ranked blogs have found a common goal to better serve the potential customers and business by means of presenting topics relevant to their industry. In constructing a topic in business communication, different resources were drawn on, including theoretical knowledge, personal experiences; however, the process type of those highly frequent words such as 'using', and 'making' is basically actional, and they are colligated with a variety of meanings and very wide contexts.

Secondly, the construction of CB register was interrelated with the blog writers' decision-making based on their perceptions on related issues as well as the contextual issues. For instance, in the process of constructing topics in CB texts, whether it should be professional or not professional was somewhat determined by the industry they belong to; however, a conversational like tone can be more easily controlled. Similarly, it was also their own decision on whether or not to put some personal or product-related topics in one website, depending on their perception of the function of the online blogs. Take blog web site **Delta** as an example; both texts from text type 3 (interactive storytelling) and text type 5 (persuasion) are included in this site. In '*Behind-the-Wing: Passenger Stories*', a personal story was told, while in the text named 'Y*our Trip (Extras) – Your Choice*', the audiences were persuaded directly by saying '*Now you can give that Sky Miles account some added oomph by purchasing some extra miles for your account*'.

Thirdly, this study has also presented complexities the blog writers and their readers were faced with. For instance, in some blog writers' minds, corporate blogs were designed for a business purpose; however, the readers might also use it for personal and social purposes. Whereas the blog writers expected the readers to become their loyal customers, at the same time, he/she acknowledged that the friendly, personal, and plainspoken interpersonal element was indispensable for the development of the online community. The blog readers were also faced with dilemmatic situations. For example, in the friendly atmosphere of the online blogs, where everyone was trying to be involved in it, the readers possibly found some incomprehensible professional knowledge or uninteresting personal stories.

Fourthly, in the analyses of the tenor and mode construction, a central theme presented was the concerns about interpersonal relations. With regard to tenor, the blog writers were careful in making linguistic choices. They tried varied means such as conversational like tone to create mutual solidarity. As for mode variation, CB is basically a written discourse, which may use a wide range of a complex structure rather than terminologies, for the purpose of creating a more open and comfortable atmosphere. Overall, after looking at these multiple dimensions, a conclusion can be drawn that what is highlighted in constructing the interpersonal relations of

corporate blogs suggests that corporate blogs, as an online register as well as a component of social media, is very much like exchanging information in a face-to-face setting.

Chapter 7: Conclusions

This chapter presents a summary of findings in this research project (Section 7.1), followed by a discussion of the implications from the study (Section 7.2), and some limitations of the study (Section 7.3). Finally, issues that deserve further studies are discussed in Section 7.4.

7.1 Summary of findings

Before summarizing the major findings in this study, I would like to revisit the research objectives presented in Section 1.4. This research has set three main objectives: 1) to identify the principal dimensions of linguistic variation and the primary text types in corporate blogs; 2) to reveal the relationship between CB text types and industries; 3) and identify salient linguistic differences across text types in corporate blogs in the configuration of some components of the context of situation. Taken together, the ultimate goal of this research project is to provide a comprehensive linguistic description of variations among top-ranked corporate blogs, on the basis of a corpus of texts that represents corporate blogs, and to consider the full set of lexico-grammatical features that are potentially relevant for identifying the variations. In order to achieve these goals, a series of analyses were conducted in light of the theoretical framework involving both MD analysis and SFL. The findings of these analyses suggest that the corporate blog represents a rather complex genre being embedded in a highly varied online business setting. Under the umbrella term 'register', several aspects of variation were addressed and proved to be inseparable from each other, including the variation according to text types, the industries companies belong to, and the context of situation. The findings also indicate that Biber's MD analysis was successfully combined with Halliday's SFL approach which made possible a linking between different aspects in the semiotic dimension of "cline of instantiation" (Halliday & Matthiessen, 2004, p. 27), providing us a wider and a more comprehensive picture of the functionally linguistic variation in corporate blogs.

By using an MD approach, six factors/dimensions and six text types were identified and then interpreted. Six factors/dimensions include: (1) personal involvement narration, (2) informational production, (3) interactive persuasion, (4) abstract/impersonal style, (5)

evaluative stance, and (6) general information vs. technical information. This result suggests that there is continuous variation along multiple dimensions of the language of corporate blogs. On all dimensions, the corporate blogs were found to cover a broad spectrum of registers, with a series of co-occurring linguistic features reflecting a wide range of communicative functions. Some dimensions comprise several sets of functions related to the primary purpose of communication, such as Dimension 1 which combines functions related to personal involvement and narrative discourse, and Dimension 3 which combines functions related to an interactive style and argumentative/persuasive discourse. With regard to Dimension 1 (Personal involvement narration), the co-occurring functions were actually distinct dimensions identified in previous studies, i.e., Dimension 1 (personal, involved/stance-focused narration) in Biber and Kurjian (2006), Dimension 1 (Involved production) and Dimension 2 (narrative discourse) in Biber (1988). Similarly, as for Dimension 3 (interactive persuasion), the co-occurring functions could be found in two distinct dimensions in Biber (1988), i.e., Dimension 1 (Involved production) and Dimension 4 (Overt expression of persuasion). As for the other dimensions, including Dimension 2, Dimension 4, and Dimension 5, there is a single communicative function, i.e., informational production, abstract/impersonal style, and evaluative stance, associated with each set of linguistic features. Dimension 6 is special in that it consists of several elements, being either general/technical, formal/informal, or spoken/written, constituting a multi-layer dimension.

As for the relationship between these dimensions in corporate blogs and written/spoken registers, they are interweaved with each other: some dimensions could identify sharp distinctions between spoken and written registers, some are the mixture of those communicative functions that may or may not distinguish spoken and written registers. Following the discussion of Biber (1988), it could be observed that the communicative function of involved production in Dimension 1 (Personal involvement narration) and Dimension 3 (Interactive persuasion) marks affective, interactional and highly involved text production. Besides, written discourse is often ‘informational’ (Dimension2), ‘impersonal’ (Dimension 4), and ‘situational-dependent’ (Dimension 5). Some communicative functions, including narrative discourse, and persuasion, do not have a systematic relationship to spoken and written

language; however, they combine some features of spoken registers, and form a rather hybrid dimension, e.g., Dimension 1 (Personal involvement narration), and Dimension 4 (Interactive persuasion).

With the dimensions being further used as indicators, a following-up cluster analysis was conducted and six text types were identified: (1) descriptive discourse, (2) technical conversation, (3) interactive storytelling, (4) evaluative information report, (5) persuasion and (6) interactive instruction. This result suggests that corporate blog is a hybrid genre, consisting of various text types. These text types do not fit into any existing text types, like narration, argumentation, exposition, etc. (Werlich, 1976); rather, they are hybrid in nature, with several language features and functions co-occurring in one text type. The variation in CB text types is not in chaos, but in the continuum: an oral-literate continuum and a narrativity continuum. On the narrativity continuum, text type 3 (interactive storytelling) consists of some core narrative texts with some core narrativity features, focusing on telling a story. Some other text types, including text type 1 (descriptive discourse), text type 2 (technical conversation), text type 4 (evaluative information report), and text type 5 (persuasion), do not primarily focus on telling a story, but include narration as a secondary or simultaneous objective alongside primary objectives and underlying functions such as scientific exposition, persuasiveness, information presentation or interpersonal interaction. On the oral-literate continuum, the basic communicative functions from various text types constitute a paradigm from involved production representing spoken register to some written discourse features, like 'informational' (from text type 6), 'impersonal' (from Cluster 4), and 'situational-dependent' (from text type 3 and text type 4). When different communicative purposes and CB text types coordinate with each other, some more general strategies and organizational purposes could be observed. Among the varied communicative functions in corporate blogs, it is the marked style of being highly interactive and informational that makes corporate blogs keep a friendly atmosphere, and a standard blog voice. Some other functions, such as informational production, persuasion, abstract/impersonal style may help develop new product or even sell some products directly to a wide range of audiences, and help solve problems by giving instructions, and explanations, etc. The communicative function of storytelling may play an essential role in building image

and establishing community between readers and writers.

Taking text type 2 (technical conversation) as an example, which takes up the largest percentage of the whole data, texts from this text type focus on technical issues, and the language employed in this text type functions to close the social distance between readers and writers, by means of 'talking' to their professional rather than amateur readers in a conversational tone. Thus, it can be concluded that, apart from being characterized by a conversational tone when presenting information, the corporate blog is fundamentally industry-based. That means the knowledge or any information published by a corporate blog is possibly largely confined to the industry it is from. In view of this, it was decided that the relationship between industries and CB text types should be explored, the purpose of which is to provide more fully the interpretation of CB text types. Before this exploration, the corporate blogs in CBC were divided into 8 industries, on the basis of the background information of two classification schemes, as well as the nature of corporate blogs. The two systems for the industrial classification of corporate blogs include International Standard Industrial Classification (ISIC) of All Economic Activities (see Appendix A3), and the Global Industry Classification Standard (GICS) system (see Appendix A4). And the 8 industries are: service industry, wholesale and retail trade, manufacturing, activities of extraterritorial organizations and bodies, publishing activities, e-commerce and online social networking service, information and communication, information technology.

The main conclusion from the investigation of the relationship between CB text types and industries is there is a close relationship between them. They are text type 1 (descriptive discourse) and the publishing industry, text type 3 (interactive storytelling) and the wholesale and retail trade industry, text type 4 (evaluative information report) and the industry of Information and communication & publishing industry, text type 5 (persuasion) and the industry of e-commerce and online social networking service, text type 6 (interactive instruction) and the whole sale and retail trade industry. In contrast, CB texts from text type 2 (technical conversation) are relatively proportionally distributed across different industries. The close relationship between certain text types and industries implies that some important strategies

were possibly adopted for creating and publishing a successful corporate blog. This could be clearly exhibited if the co-occurring language features in text types are associated with some industry attributes. In particular, for the corporate blogs from the publishing industry, they are responsible for making statements of facts, and linguistically, they are dominated by texts from text type 1 (descriptive discourse), in which some informational-focused language features with an impersonal-style frequently occurred. The close relationship between text type 3 (interactive storytelling)/text type 6 (interactive instruction) and the whole sale and retail trade industry shows that the sales strategy in corporate blogs is against the traditional way of using some persuasive language; rather, storytelling and giving instructions are employed in their branding or selling. In contrast, it is in the industry of e-commerce and online social networking service that the persuasive language is more frequently used. The frequent use of persuasive language is a strategy for coping with the nature of e-commerce websites. Since e-commerce is the sale of products and services over the internet, the use of persuasive language may help convince the visitors of the website that the products offered are right for them, although these potential customers can't see or touch the products that are not physically in front of them. In this way, the e-commerce site visitors could possibly be converted into loyal customers.

For the corporate blogs from the industry of information technology, they are most closely related to text type 2 (technical conversation), which means linguistically they inevitably include many technical terms or terminologies. The corporate blog from this industry may cover a broad range of topics concerning information technology, such as the use of electronic computers, network administration, server maintenance, and computer software, etc., which is usually not easy to understood by the general public. In spite of this, the use of an 'informal and friendly talking' strategy with more human voice may ease the tension, and make possible the knowledge sharing between firms and customers, since the potential customers are possibly experts or a layman. All these attributes and the correlations between certain text types and industries indicate the variability in corporate blogs. For corporate blogs from various industries, they have their own type of role. Each company has a custom approach based on the products it sells, its customer demographic and service approach, and the overall

company culture. Yet, most corporate blogs prefer to use some indirect brand marketing initiatives as the priorities, such as giving instruction or providing some professional information in a casual and interactive way, telling stories in blogs from sales industry, etc. Even in a blog created by a particular company, it consists of CB texts from different text types, which means a variety of language features and strategies are used by a company. A company may not only need to tell stories, but also need to provide customer service or some help directly; they may also need to introduce some background and product information. In this way, different purposes of corporate blogs, including building the loyalty of blog visitors, as well as selling goods directly to them, are achieved.

Based on the investigation of different CB text types across industries, a conclusion can be drawn: as a means of social media, the corporate blog gives firms a way to communicate with peers, customers and potential consumers. Taken together, different communicative purposes as well as the strategic decisions in different text types may constitute a widely implemented model, in which three broad categories of CB are included: sales, customer relationship management and branding. In this model, there is always a tension created and adjusted dynamically to balance the social distance between authors and readers. The authors have always tried to build the 'brand' in a relaxed and conversational way. Branding can be called the business identity, a way to define the company's business to the external audiences. It is not just a logo or about how their business is recognised; in corporate blogs, branding is built and thought of as an individual person whose character is made up of beliefs, values and purposes. No matter what information are exchanged, e.g., the storytelling, the instruction, professional knowledge, they represent different personalities behaving in different situations. This helps reinforce the business' character and clarify its offering so that customers are aware of what to expect from the product or service. At the same time 'speaking' to the audiences in a consistently conversational and friendly way may help build a long-term relationships with the customers. Overall, the three general purposes of CB which include sales, customer relationship management and branding are able to be achieved simultaneously, and sometimes by the same communicative strategies.

Following the framework for this research which combines MD analysis and the SFL-based approach, the differences of several linguistic features (pronouns, modal verbs, lexical density and grammatical complexity) across the 6 text types were further identified. The results of this study show that there is a significant difference of the use of these language features across six text types. As for the most common verbs across six text types, ***said, using, found, using, need, want*** are the most common verbs in each of the six text types respectively; among the fifteen most common verbs in CBC, the verb ***make*** is the only one that takes all the tenses (***make, makes, made, has/have made, making***) and patterns, thus constituting a great proportion of occurrences. In terms of the use of personal pronouns, the results correspond to the personal principle claimed by Maslansky et al. (2010): the personal pronouns '*I*', '*we*' and '*you*' are dominantly used in CB, in comparison with other types. Besides, there is a clear-cut distinction of the use of '*you*' in text type 5 (persuasion), which shows that the blog writer aims to send messages, with the hope of bonding with his audiences, as well as gaining trust from them. When it comes to the modal verbs, there is a contrastive use of possibility modals and necessity modals, particularly in text type 5 (persuasion). This shows that blog authors have been trying to avoid the use of languages of commanding, or creating obligations when they try to persuade their readers to accept a new product. Rather, they hope to convince their audiences by means of making suggestion, and presenting logical possibility or ability to be acquired. Besides, most CB text types are close to each other in terms of their lexical density and grammatical complexity, apart from text type 1 (descriptive discourse). Text 1 has the smallest LD, the longest MLS, MLT, and the most complex CN/C and CN/T. Thus, it could be seen as the most spoken text type, which is less lexically dense and more complex grammatically.

Taken together, field, tenor and mode related language features in corporate blogs demonstrate a dynamic nature in it: centring around an interpersonal function, the online blogs in a business setting was used for different purposes, e.g., promoting products, delivering latest news, telling stories. In spite of this, they are closely related to the industry they belong to. When a blog is constrained to a particular industry, some tensions possibly emerge, including those between the interpersonal interaction and factual objectivity in a text, and

those of social distance between authors and readers. In order to ease the tension, the language of corporate blogs is varied from highly formal technical language to very interactive language. Therefore, when delivering different information, language style is changed for establishing a tone as friendly as possible.

7.2 Implications of the study

In this research project, the multi-dimensional analysis and the SFL-based approach for the identification of variation in corporate blogs have made some theoretical, methodological and empirical implications, as presented below.

Theoretically, the combination of multi-dimensional analysis and the SFL in this study advances the research on register variation. In practice, multi-dimensional studies on any other registers or genres have not been integrated with any other methodological studies. In this research, the precedence of multi-dimensional analysis before the SFL-based approach constitutes a framework encompassing both bottom-up and top-down approaches that are assumed to provide a more accurate, more detailed, and more thorough template about communicative practices of corporate blogs. Guided by this framework, different aspects of variation in corporate blogs have been investigated and evaluated, including the variation according to text types, variation according to industries, and variation according to the context of situation. These different aspects of variation are inseparable from each other; they are actually some closely related patterns on “the cline of instantiation” (Halliday & Matthiessen, 2004, p. 28). When identifying a text type, we “move along the cline of instantiation away from the text pole towards the system pole” (Halliday & Matthiessen, 2004, p. 27). When identifying the context of situation across text types, there is not only a moving alone the cline of instantiation, but also a comparative perspective looking into these patterns. The contribution of this research project is it provides some empirical support, which makes possible the move along the spectrum from theory to practice. Thanks to the empirical analyses in this research, the variations of text types, industries, and context of situation are connected to each other. The variation of text types can be detected among clusters of texts, and in each clusters, some functional co-occurrences of linguistic features are shared. On this basis, looking into how the

variation of industries cuts across the variation of text types may help reveal the linking between business communication in a particular industry and certain text types characterized by some co-occurring linguistic features; in this way, texts with similar text-internal features are related with certain business function that is roughly similar across industry groupings, and a more comprehensive interpretation to CB text types can be made. Furthermore, the exploration of three metafunctions from context of situation across different CB text types provides more evidence as for to what extent each text type can be considered as a semantic variety in the linguistic system, and realizes its own configurations of lexico-grammatical features. In particular, different metafunctions are primarily dominant in different text types. The ideational metafunction is primary in the storytelling, descriptive and instructive discourse type, which is action-oriented. In other text types, such as text type 2 (technical conversation), text type 4 (evaluative information report), and text type 5 (persuasion), the writers tend to intrude into the minds of their readers through explaining, persuading, etc., and the interpersonal metafunction is more likely to be the label of these text types. Basically, the corporate blog is influenced by three metafunctions simultaneously that are woven together into the same linguistic unit. To sum up, the analyses conducted in this research concern different aspects of variation.

In comparison with the representative work by Biber (1988), which is based on the multi-genre corpora, the current research focused on a single genre of corporate blogs. In spite of being narrowed down to a single genre, the current research also gets a broad communicative functions we see in the analyses of multi-genre corpora (Biber, 1988), as well as some web-based corpora (Biber & Kurjian, 2006). The text types identified could be characterized as either a narrative or a peripheral narrative style. As shown in Chapter Five, five out of six text types were identified as capturing subtle characteristics of telling a story or primarily focusing on telling a story, which means including narration as a secondary or simultaneous objective alongside primary objectives and underlying functions such as scientific exposition, persuasiveness, information presentation or interpersonal interaction. Even in text type 6 (Instruction), the most ‘non-narrative’ text type, it is mainly concerned with telling someone what to do and how to do something through an interactive, narrative and somewhat

persuasive tone. Therefore, the corporate blog is informational and interactional in nature. This also accounts for the metafunctional orientation to different text types elaborated above: the ideational metafunction is primary in action-oriented text types, e.g., storytelling (narration), instruction, etc., while the interpersonal metafunciton is primary in some explanatory or argumentative/persuasive text types.

No matter how wide communicative functions corporate blogs have, or how hybrid the corporate blogs are, what makes it a distinctive genre is its industry-oriented attribute. Based on an added procedure in MD which focuses on the relationship between CB text types and industry, it is found that CB texts from a certain text type are mainly from an industry type, such as text type 2 (technical conversation) and the industry of information technology, text type 1 (descriptive discourse) and the publishing industry, text type 3 (interactive storytelling)/text type 6 (interactive instruction) and the whole sale and retail trade industry, text type 5 (persuasion) and the industry of e-commerce and online social networking service. Adding this procedure to a traditional MD analysis stems from the consideration that the more situational and contextual factors are added, the better; in this way, a more valid and more comprehensive interpretation of CB text types can be made.

On this basis, a further step I made is exploring the variation according to context of situation in corporate blogs, the purpose of which is discovering the wealth of meanings in corporate blogs, as well as how meanings can be distinguished across different CB text types. The SFL metafunction model has provided a perspective to view language as a complex of choices between mutually exclusive options. I have explored some elements generated by ideational, interpersonal and textual metafunction, which are distinct from each other. The exploration on the most common verbs and their process types from ideational metafunction, the attitudinal elements from the interpersonal metafunction, and the issue of the spoken/written axis from the textual metafunction helps us understand further the preferred mode of communication, how different attitudes are reflected in different text types, and how complex or intricate the language is when delivering their message to the readers, etc. Taken together, the analysis of prominent metafunctional elements produced specific contexts that acted as a social

happening in corporate blogs. As a further analytical work to the previous MD analysis, I consider that this SFL-based analysis is plausible, in that the SFL methodology is enriched with MD analysis in some way, as well as making possible a shift from the introspection of tangible language features to intangible form-function relations. In other words, with the language features as the starting point for this research, the meaning-making at a descriptive level can be accessed ultimately. In this way, the gap between different patterns, e.g. text types, texts, context of situation, at either micro- or macro-levels are bridged. The profound semiotic processes involve lexico-grammatical and textual choices up to discoursal ones as a continuum system, in which a minimum balance is kept. In a broad sense, this framework can be employed effectively in various types of language use.

Overall, this investigation into the language of corporate blogs contributes to the existing field of knowledge in two ways: Firstly, it develops the methodology used in corpus investigation of language variation. The contributions made to methodology by this research include the role of CB discourse as an example of corpus-based investigation; and the application of MD analysis and corpus-based study on lexical and grammatical features of CB. Taken together, these aspects pave the way for further research into corporate blog and other forms of electronic communication and, more generally, for researchers engaging in corpus-based investigations of other language varieties. Secondly, it adds greatly to a description of corporate blog as a language variety in its own right. The contribution that this research makes towards the description of corporate blog as a language variety include different text types identified in CB discourse, and some linguistic features realized in the context of situation. This highlights the fact that corporate blog cannot be regarded as a simple discourse; rather, it varies according to text types and context of situation.

7.3 Limitations of the study

In the current research, the combination of MD analysis and an SFL-based approach is innovative; however, this study has tended to provide an in-depth analysis; however, the limitation is it lacks breadth and intensity. It is mainly due to nature of corporate blogs, which is a domain-specific type of blog from a domain-specific genre, that it differs from those general

spoken/written language representing more general disciplines, and thus it lacks representation of some disciplines in the analysis. This impedes the interpretation of the results generated by this research. And the drawbacks can be seen in various stages of the analysis.

In Factor analysis, a component part of MD analysis, and one of the major statistical analyses to reduce a large set of variables into a smaller set, many decisions have to be made. One of them is to decide how many variables (here linguistic features) can be included. The decision of this issue in the current research has adopted some principles, such as marker variables, which means the variables included in a factor analysis should also be correlated and reliable, the availability for tagging these features by using some reliable computer programs. It has also been justified that most factors identified in this research is represented by a moderate amount of variables in the analysis. However, in some factors, the problem is they do not consist of a minimum of five variables (Biber, 1988). In this view, I suppose there are possibly some more accurate decisions ahead of me. Based on these principles, it is inevitable to encounter a problem that with limited functions of computer programs I use, some key features might be excluded. Besides, even if the number of variables included in this research has been proved to be accurate and big enough, the features as well as the particular communicative functions singled out are not comparable to those from more general corpora.

There are also many uncertain issues in cluster analysis, and even the classification of industries. The determination of the number of clusters is a difficult issue we often encountered in cluster analysis. Although the *pseudo-F* test (Caliński & Harabasz, 1974) which is assumed to have a better performance is employed, it should be admitted that this issue is still an unsolved problem of cluster analysis (Everitt, 1979). Many researchers consider there are no completely satisfactory methods yet for the determination of the number of clusters for any type of cluster analysis (Bock, 1985; Hartigan, 1985). As for the classification industries, there are many different standards released by different groups from different countries. Although what has been adopted for the industry classification in the current research are prestigious and widely used, the difficulty of categorizing a corporate blog into an industry is the company

publishing that blog might belong to several different industries, if viewed from different scopes in business. When encountering these problems, I take into account the firm's primary business activities, and categorize it into a major industry sector.

7.4 Directions for future studies

This research has made contribution to the description of corporate blogs, providing a rich and comprehensive analysis of language use in an online and business setting. Considering that there are still many questions awaiting us to answer, future research is called for to expand our understanding of how corporate blogs, as well as other types of CMC are constructed. The corporate blog is such a complex undertaking and meaning making process that a variety of perspectives can be focused on.

At first, in future studies on corporate blogs, some quantitative techniques can be adopted for a more sensible selection of variables. And simultaneously, some superior statistical tools for some additional features, such as those from grammatical categories, can be used. This could possibly uncover register variations more precisely and thoroughly. In future studies, data resources can be broadened, so that the studies on the use of blog language can be enriched. The data could be collected on the basis of different criteria, e.g., by different population groups, from a gender point of view, by a particular person, or by different types of companies, etc.

There is also a clear need for the research on blogs from a broader range of disciplines. It will be interesting to see in other disciplines what kind of text types can be identified and how context of situation are realized linguistically. It is also of great necessity to look into blogs of other languages. Cross-linguistic comparisons will also benefit those who have non-English backgrounds, in the way they find the differences across languages and learn to bridge such gaps, so that they are able to present their information more successfully. Without this body of knowledge, there is a risk that we are not able to have a deeper insight into the cultural influence on blog language.

In a broad sense, a general feature of CMC has been identified, as also found in corporate blogs: there are some short cuts, e.g., abbreviations, acronyms, and unique spellings, which were said to be motivated by the technology in use and marking CMC language. In this view, it would be useful to focus on these issues and carry out some further research. CMC is such a broad term which includes any means of communication by using a computer or on computer net-work. Then, blogs, or other CMC types, such as wikis, emails can be compared to each other in the larger CMC domain.

In relation to my findings, it appears that SFL is such a comprehensive framework encompassing various components, that, in the future, more linguistic features can be chosen and analysed by using this framework where features can be mapped onto the Halliday's metafunctions of language. Further research needs to be done in order to substantiate these assumptions. Besides, the drawbacks of both MD analysis and SFL theory have been proposed by some researchers; it is also hoped that some improvements can be made technologically and theoretically.

7.5 Summary

This research demonstrates that the linking of multi-dimensional analysis and SFL can enhance our understanding of how a representative sample of corporate blogs is varied. Different dimensions are realised by different sets of co-occurring patterns, and text types are identified to be significantly different from one another and exhibit a hybrid nature. On these basis, whether in the exploration of the relationship between CB text types and industries, or the investigation of those language features associated with context of situation (pronouns, modal verbs, lexical density, or the grammatical complexity), it is not only the varied communicative purposes but also the broad categories of CB that could be identified: sales, customer relationship management and branding. It has also been perceived that a balanced strategy is widely used by the firms: blog writers tend to be very careful in delivering the information so as to maintain a warm and friendly online atmosphere. Overall, it has been proved quantitatively that both a 'personal principle' and a 'plainspoken principle', as a foundation of the language of trust, have been followed in the CBC, which represents

top-ranked successful corporate blogs. Besides, the power of the analytical framework offers the opportunity to empirically answer the research questions. The findings from this research contribute to our understanding of the variation of corporate blogs.

REFERENCES

Aldenderfer, M. S., & Blashfield, R. (1984). *Cluster analysis*. Newbury Park, CA: Sage Publications.

Aleknavičiūtė, T. (2009). Linguistic Features of Political Blogs in English and Lithuanian. *Žmogus ir žodis, 1*(11), 13-20.

Altenberg, B. (1984). Causal linking in spoken and written English. *Studia linguistica, 38*(1), 20-69.

Alvesson, M., & Karreman, D. (2011). *Qualitative research and theory development: Mystery as method*. London: Sage Publications.

Andersen, P. B. (2002). Dynamic semiotics. *Semiotica, 139*(1/4), 161-210.

Anderson, S. R. (1971). On the role of deep structure in semantic interpretation. *Foundations of language, 7*(3), 387-396.

Anis, J. (2007). Neography: Unconventional spelling in French SMS text messages. In B. D. S. C. Herring (Ed.), *The multilingual internet: Language, culture, and communication online* (pp. 87-115).

Asencion-Delaney, Y., & Collentine, J. (2011). A Multidimensional Analysis of a Written L2 Spanish Corpus. *Applied linguistics, 32*(3), 299-322(224).

Aston, G., & Burnard, L. (1998). *The BNC handbook: exploring the British National Corpus with SARA*. Edinburgh: Edinburgh University Press.

Atkinson, D. (1992). The evolution of medical research writing from 1735 to 1985: the case of the Edinburgh Medical Journal. *Applied Linguistics, 13*(4), 337-374.

Atkinson, D. (1993). *A historical discourse analysis of scientific research writing from 1675 to 1975: The case of the philosophical transactions of the Royal Society of London.* (Unpublished PhD dissertation), University of Southern California.

Bahnisch, M. (2006). The political uses of blogs. In e. Axel Bruns & Joanne Jacobs (Ed.), *Uses of blogs* (pp. 139-149). New York: Peter Lang.

Banks, D. (2008). *The Development of Scientific Writing. Linguistic features and historical context*. London: Equinox.

Bar-Ilan, J. (2004). *An outsider's view on topic-oriented blogging.* Paper presented at the Proceedings of the 13th international World Wide Web conference on Alternate track papers & posters.

Barnes, N. G., & Lescault, A. M. (2011). Social media adoption soars as higher-ed experiments and reevaluates its use of new communications tools. *Center for Marketing Research. University of Massachusetts Dartmouth, North Dartmouth, MA*.

Barnes, S. B. (2003). *Computer-mediated communication: Human-to-human communication across the Internet*. Boston: Allyn and Bacon

Baron, N. S. (1998). Letters by phone or speech by other means: The linguistics of email. *Language & Communication, 18*(2), 133-170.

Barra, M. (2008). Global Industry Classification Standard. List of code and description, Microsoft Excel format. http://www.mscibarra.com/resources/xls/GICSmap2008.xls.

Barthes, R. (1966). Introduction à l'analyse structurale des récits. *Communications, 8*(1), 1-27.

Baym, N. K. (1995). The emergence of community in computer-mediated communication. *In S. G. Jones (Ed.), Cybersociety (pp. 138-163). Thousand Oaks, CA: Sage.*

Beaugrande, R. d. (1993). 'Register'in Discourse Studies: a concept in Search of a Theory *Register analysis. Theory and practice* (pp. 7-25).

Bhatia, V. K. (1993). *Analysing Genre: language use in professional settings*. London: Longman.

Biber, D. (1985). Investigating macroscopic textual variation through multifeature/multidimensional

analyses. *Linguistics, 23*(2), 337-360.

Biber, D. (1986). Spoken and written textual dimensions in English: Resolving the contradictory findings. *Language, 62*(2), 384-414.

Biber, D. (1988). *Variation across speech and writing*. Cambridge: CUP.

Biber, D. (1989). A typology of English texts. *Linguistics, 27*(1), 3-44.

Biber, D. (1993). Representativeness in corpus design. *Literary and linguistic computing, 8*(4), 243-257.

Biber, D. (1994). An analytical framework for register studies. In Biber, D. & Finegan, E. (eds.), Sociolinguistic Perspectives on Register. Oxford University Press. (pp. 31-56).

Biber, D. (1995). *Dimensions of register variation: A cross-linguistic comparison*. Cambridge: Cambridge University Press.

Biber, D. (2001). Dimensions of variation among eighteenth-century speech-based and written registers. *Variation in English: multi-dimensional studies*, 200-214.

Biber, D. (2004a). Conversation text types: A multi-dimensional analysis. *Le poids des mots: Actes de JADT*, 15-34.

Biber, D. (2004b). Historical patterns for the grammatical marking of stance: A cross-register comparison. *Journal of Historical Pragmatics, 5*(1), 107-136.

Biber, D. (2004c). Modal use across registers and time. *Topic in English linguistics, 45*, 189-216.

Biber, D. (2006). *University language: A corpus-based study of spoken and written registers*: John Benjamins Publishing.

Biber, D. (2012). Register as a predictor of linguistic variation. *Corpus Linguistics and Linguistic Theory, 8*(1), 9-37.

Biber, D., & Conrad, S. (2003). 9 Register Variation: A Corpus Approach. In D. Schiffrin, D. Tannen, & H. E.. Hamilton (Eds.), Handbook of discourse analysis. (pp. 175-196).

Biber, D., & Conrad, S. (2009). *Register, genre, and style*. Cambridge: Cambridge University Press.

Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus linguistics: Investigating language structure and use*. Cambridge: Cambridge University Press.

Biber, D., Conrad, S., Reppen, R., Byrd, P., & Helt, M. (2002). Speaking and writing in the university: A multidimensional comparison. *Tesol Quarterly, 36*(1), 9-48.

Biber, D., Conrad, S., Reppen, R., Byrd, P., & Helt, M. (2003a). The authors respond: strengths and goals of multidimensional analysis. *tesol Quarterly, 37*(1), 151-155.

Biber, D., Conrad, S., Reppen, R., Byrd, P., & Helt, M. (2003b). Strengths and Goals of Multidimensional Analysis. *Tesol Quarterly, 37*(1), 151-156.

Biber, D., Davies, M., Jones, J. K., & Tracy-Ventura, N. (2006). Spoken and written register variation in Spanish: A multi-dimensional analysis. *Corpora, 1*(1), 1-37.

Biber, D., & Finegan, E. (1989a). Drift and the evolution of English style: A history of three genres. *Language*, 487-517.

Biber, D., & Finegan, E. (1989b). Styles of stance in English: Lexical and grammatical marking of evidentiality and affect. *Text-Interdisciplinary Journal for the Study of Discourse, 9*(1), 93-124.

Biber, D., & Finegan, E. (1992). The linguistic evolution of five written and speech-based English genres from the 17th to the 20th centuries. In O. I. In M. Rissanen, T. Nevalainen, and I. Taavitsainen (eds.). History of Englishes: New methods and interpretations in historical linguistics. (Ed.), (pp. 688-704). Berlin: Mouton de Gruyter.

Biber, D., & Finegan, E. (1994a). Intra-textual variation within medical research articles. *Corpus-based research into language, 12*, 201-222.

Biber, D., & Finegan, E. (1994b). *Sociolinguistic perspectives on register*. Oxford: Oxford University Press.

Biber, D., & Finegan, E. (2001). Diachronic relations among speech-based and written registers in English. *Variation in English: multi-dimensional studies*, 66-83.

Biber, D., Finegan, E., & Atkinson, D. (1994). ARCHER and its challenges: compiling and exploring a representative corpus of historical English registers. In: Fries, U., Tottie, G. & Schneider, P. (eds) Creating and Using English Language Corpora. Amsterdam: Rodopi, 1-14.

Biber, D., Finegan, E., Atkinson, D., Beck, A., Burges, D., & Burges, J. (1994). The design and analysis of the ARCHER Corpus: a progress report. *Kytö et al.(eds.)*, 3-6.

Biber, D., Gray, B., & Poonpon, K. (2011). Should we use characteristics of conversation to measure grammatical complexity in L2 writing development? *Tesol Quarterly, 45*(1), 5-35.

Biber, D., & Hared, M. (1992). Dimensions of register variation in Somali. *Language Variation and Change, 4*(01), 41-75.

Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). Grammar of spoken and written English. *Harlow: Longman Pearson*.

Biber, D., & Kurjian, J. (2006). Towards a taxonomy of web registers and text types: a multidimensional analysis. *Language and Computers, 59*(1), 109-131.

Blood, R. (2002). *The weblog handbook: Practical advice on creating and maintaining your blog*: Basic Books.

Bloor, T., & Bloor, M. (1995). *The Functional Analysis of English: A Hallidayan Approach*. London: Arnold.

Bock, H.-H. (1985). On some significance tests in cluster analysis. *Journal of classification, 2*(1), 77-108.

Bortree, D. S. (2005). Presentation of self on the web: An ethnographic study of teenage girls' weblogs. *Education, Communication & Information, 5*(1), 25-39.

Bowker, N., & Tuffin, K. (2004). Using the online medium for discursive research about people with disabilities. *Social Science Computer Review, 22*(2), 228-241.

Brady, M. (2005). Blogging, personal participation in public knowledge-building on the web. *Chimera Working Paper, 2*(11), 4-16.

Brecht, F., Cudreasova, A., & Zhou, J. (2010). Corporate blogging today–usage and characteristics *Sustainable e-Business Management* (pp. 87-99). Springer Berlin Heidelberg.

Brown, P., & Fraser, C. (1979). Speech as a marker of situation *Social markers in speech* (pp. 33-62). Cambridge: Cambridge University Press.

Brown, P., & Levinson, S. C. (1987). *Politeness: Some universals in language usage* (Vol. 4). Cambridge: Cambridge University Press.

Brown, S. P. (2008). Business Processes and Business Functions: a new way of looking at employment. *Monthly Lab. Rev., 131*, 51.

Browne, M. W. (1968). A comparison of factor analytic techniques. *Psychometrika, 33*(3), 267-334.

Brownstein, E., & Klein, R. (2006). Blogs: Applications in science education. *Journal of College Science Teaching, 35*(6), 18.

Buckingham, D., & Willett, R. (2013). *Digital generations: Children, young people, and the new media*. New York: Routledge.

Burnard, L. (1995). *Users Reference Guide British National Corpus Version 1.0*. Oxford: Oxford University Computing Services.

Burton, D. (1980). *Dialogue and Discourse: A Sociolinguistic Approach to Modern Drama Dialogue and Naturaely Occurring Conversation*. London: Routledge and Kegan Paul.

Butt, D., Fahey, R., Spinks, S., & Yallop, C. (2000). Using Functional Grammar: An Explorer's Guide. National Centre for English Language Teaching and Research. *Macquarie University, Sydney*.

Butt, D. G. (2004). Parameters of Context: On establishing the similarities and differences between social processes. *Unpublished manuscript, Sydney*.

Caffarel, A., Martin, J. R., & Matthiessen, C. M. (2004). *Language typology: A functional perspective*. Amsterdam: John Benjamins Publishing.

Caliński, T., & Harabasz, J. (1974). A dendrite method for cluster analysis. *Communications in Statistics-theory and Methods, 3*(1), 1-27.

Cattell, R. (1978). *The scientific use of factor analysis in behavioral and life sciences*. New York: Plenum.

Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research, 1*(2), 245-276.

Celce-Murcia, M. (1998). Discourse analysis and grammar instruction. *In D. D. Oaks (Ed.), Linguistics at work: A reader of applications (pp. 687-704). Fort Worth, TX: Harcourt Brace.*

Chafe, W. (1982). Integration and involvement in speaking, writing, and oral literature). In D. TANNEN (Ed.): Spoken and Written Language: Exploring Orality and Literacy. Norwood, NJ: Ablex Publishing Corporation.

Chafe, W. (1985). Linguistic Differences Produced by Differences Between Speaking and Writing. *In D.R.Olson, N. Torrance, & A. Hildyary (Eds.), Literacy, language, and learning (pp. 105-123).*

Chafe, W., & Danielewicz, J. (1987). *Properties of spoken and written language*. San Diego: Academic Press.

Chafe, W., & Tannen, D. (1987). The relation between written and spoken language. *Annual Review of Anthropology*, 383-407.

Chafe, W. L. (1976). Givenness, contrastiveness, definiteness, subjects, topics, and points of view. *In C. N. Li (ed.) Subject and Topic. New York: Academic Press. 25-55.*

Charles, M. (2006). Phraseological patterns in reporting clauses used in citation: A corpus-based study of theses in two disciplines. *English for Specific Purposes, 25*(3), 310-331.

Charman, S. (2006). Blogs in business: Using blogs behind the firewall. *Uses of blogs, 38*, 57-67.

Chatman, S. B. (1980). *Story and discourse: Narrative structure in fiction and film*. New York: Cornell University Press.

Cheng, W. (2011). *Exploring corpus linguistics: Language in action*. London/New York: Routledge.

Chilton, P. (1978). On the Theory of Register. *Nottingham Linguistics Circular*, 7 (.2), pp. 113- 130.

Christie, F. (2004). Systemic Functional Linguistics and a theory of language in education. *Ilha do Desterro A Journal of English Language, Literatures in English and Cultural Studies*(46), 013-040.

Cloran, C. (1987). Negotiating new contexts in conversation. *Occasional Papers in Systemic Linguistics, 1*(1985), 85-110.

Close, R. A. (1978). *A reference grammar for students of English*. Harlow: Longman.

Coates, J. (1983). *The semantics of the modal auxiliaries*. London: Croomhelm.

Collot, M., & Belmore, N. (1996). A New Variety of English. *Computer-mediated communication: Linguistic, social, and cross-cultural perspectives, 39*, 13.

Comrey, A. L., & Lee, H. B. (1992). *A first course in factor analysis (2nd)*. Hillsdale: Lawrence Erlbaum Associates.

Connor-Linton, J. T. (1989). *Crosstalk: A multi-feature analysis of Soviet-American spacebridges.* University of Southern California.

Conrad, S. (2000). Will Corpus Linguistics Revolutionize Grammar Teaching in the 21st Century? *Tesol Quarterly, 34*(3), 548-560.

Conrad, S. (2001). Variation among disciplinary texts: A comparison of textbooks and journal articles in biology and history. *Variation in English: multi-dimensional studies*, 94-107.

Cook-Gumperz, J. (2006). *The social construction of literacy*. Cambridge: Cambridge University Press.

Cook, G. (1990). Trancribing infinity: Problems of context presentation. *Journal of pragmatics, 14*(1), 1-24.

Costello, A., & Osborne, J. (2011). Best practices in exploratory factor analysis: four recommendations for getting the most from your analysis. Pract Assess Res Eval 2005; 10. *URL http://pareonline. net/getvn. asp, 10*, 7.

Cowen, T. (2004). The new world of blogs. *Free-man, 54*(2), 16-18.

Coxhead, A. (1998). *An academic word list* (Vol. 18). Victoria University of Wellington: English Language Institute Occasional Publication.

Crowhurst, M. (1979). On the misinterpretation of syntactic complexity data. *English Education*, 91-97.

Crystal, D. (2001). *Language and the Internet*. Cambridge: Cambridge University Press.

Crystal, D. (2008). *Txtng: The gr8 db8*. Oxford: Oxford University Press.

Crystal, D. (2011). *Dictionary of linguistics and phonetics* (Vol. 30). New Jersey: John Wiley & Sons.

Csomay, E. (2006). Academic talk in American university classrooms: crossing the boundaries of oral-literate discourse? *Journal of English for academic purposes, 5*(2), 117-135.

Dale, R., Moisl, H., & Somers, H. (2000). *Handbook of natural language processing*. New York: Marcel Dekker.

De Dreu, C. K., Weingart, L. R., & Kwon, S. (2000). Influence of social motives on integrative negotiation: a meta-analytic review and test of two theories. *Journal of personality and social psychology, 78*(5), 889.

DeCarrico, J. S. (2000). *The structure of English: Studies in form and function for language teaching* (Vol. 1). Ann Arbor: University of Michigan Press.

December, J. (1996). Units of analysis for Internet communication. *Journal of Computer-Mediated Communication, 1*(4), 14-37.

Declerck, R. (1991). *Tense in English: its structure and use in discourse*. London: Taylor & Francis.

Divitini, M., & Morken, E. M. (2005). *Blog to support learning in the field: lessons learned from a fiasco.* Paper presented at the 5th IEEE International Conference on Advanced Learning Technologies (ICALT2005).

Dorgeloh, H., & Wanner, A. (2009). Formulaic argumentation in scientific discourse. *Formulaic language, 2*, 523-544.

Eastwood, J. (1994). *Oxford guide to English grammar*. Oxford: Oxford University Press.

Eggins, S. (1994). *An introduction to systemic functional grammar*. London: Pinter.

Eggins, S. (2004). *Introduction to systemic functional linguistics (2nd edition)*. London: Continuum.

Eggins, S., & Martin, J. R. (1997). Genres and registers of discourse. *Discourse as structure and process, 1*, 230-256.

Elgesem, D. (2002). What is special about the ethical issues in online research? *Ethics and Information Technology, 4*(3), 195-203.

Ellis, J., & Ure, J. (1969). Language varieties: register. *Encyclopedia of linguistics: Information and control*, 251-259.

Ervin-Trip, S. (1972). On sociolinguistic rules: alternation and co-occurrence. In John J. Gumperz and

Dell Hymes (eds.), Directions in sodolinguistics, 213-50. New York: Holt.

Everitt, B. S. (1979). Unresolved problems in cluster analysis. *Biometrics*, 169-181.

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological methods, 4*(3), 272.

Faigley, L., & Meyer, P. (1983). Rhetorical theory and readers' classifications of text types. *Text-Interdisciplinary Journal for the Study of Discourse, 3*(4), 305-326.

Fava, J. L., & Velicer, W. F. (1992). The effects of overextraction on factor and component analysis. *Multivariate Behavioral Research, 27*(3), 387-415.

Fawcett, R. P. (1980). Cognitive Linguistics and Social Interaction. Towards an integrated model of a systemic functional grammar and the other components of a communicating mind. *Exeter Linguistic Studies Exeter, 3*, 1-290.

Feifei, Y. (2004). Thematic Progression and Textual Coherence. *Journal of Xi'an International Studies University, 4*.

Ferguson, C. A. (1994). Dialect, register, and genre: Working assumptions about conventionalization. *Sociolinguistic perspectives on register*, 15-30.

Ferrara, K., Brunner, H., & Whittemore, G. (1991). Interactive written discourse as an emergent register. *Written communication, 8*(1), 8-34.

Fillmore, C. J. (1968). The case for case. In E. Bach & R. Harms (Eds.), Universals in Linguistic Theory (pp. 1-88). New York, NY: Holt, Rinehart and Winston.

Finegan, E., & Biber, D. (1994). Register and social dialect variation: An integrated approach. *Sociolinguistic perspectives on register*, 315-347.

Firth, J. R. (1937). *The tongues of men*. London: Watts & Co.

Firth, J. R. (1957). Papers in Linguistics 1934. *Studies in Ethnomethodology Englewood Cliffs, 244*(1), 610.

Firth, J. R. (1968). *Selected papers of JR Firth, 1952-59*. Bloomington: Indiana University Press.

Fischer, E., & Reuber, A. R. (2011). Social interaction via new social media:(How) can interactions on Twitter affect effectual thinking and behavior? *Journal of business venturing, 26*(1), 1-18.

Floyd, F. J., & Widaman, K. F. (1995). Factor analysis in the development and refinement of clinical assessment instruments. *Psychological assessment, 7*(3), 286.

Forrester, M. A. (1996). *Psychology of language: A critical introduction*. London: Sage Publications.

Foster, P., Tonkyn, A., & Wigglesworth, G. (2000). Measuring spoken language: A unit for all reasons. *Applied Linguistics, 21*(3), 354-375.

Francis, G. (1989). Thematic selection and distribution in written discourse. *Word, 40*(1-2), 201-221.

Francis, G. (1993). A corpus-driven approach to grammar: Principles, methods, and examples. In M. Baker, G. Francis & E. Tognini-Bonelli (Eds.), *Text and technology: in honour of John Sinclair* (pp. 137-156). Amsterdam: John Benjamins Publishing.

Francis, G., Hunston, S., & Manning, E. (1996). *Collins COBUILD Grammar Patterns 1: Verbs*. London: Harper Collins.

Franzosi, R. (2010). *Quantitative narrative analysis*. Los Angeles: Sage.

Fredrick, W. (2004). Your guide to corporate blogging. *Retrieved June.* 24

Gardner, S. (2008). Integrating ethnographic, multidimensional, corpus linguistic and systemic functional approaches to genre description: An illustration through university history and engineering assignments.

Geisler, C. (2002). Investigating register variation in nineteenth-century English. *Using corpora to explore*

linguistic variation, 9, 249.

Geisler, C. (2003). Gender-based variation in nineteenth-century English letter writing. *Language and Computers, 46*(1), 87-106.

Genette, G. (1983). *Narrative discourse: An essay in method*. New York: Cornell University Press.

Ghadessy, M. (1993). *Register analysis: Theory and practice*: Pinter Publishers.

Ghadessy, M. (2003). Comments on Douglas Biber, Susan Conrad, Randi Reppen, Pat Byrd, and Marie Helt's "Speaking and Writing in the University: A Multidimensional Comparison" A Reader Reacts…. *Tesol Quarterly, 37*(1), 147-150.

Ghadessy, M., Henry, A., & Roseberry, R. L. (2001). *Small corpus studies and ELT: theory and practice* (Vol. 5). Amsterdam: John Benjamins Publishing.

Gillin, P., & Schwartzman, E. (2010). *Social marketing to the business customer: Listen to your B2B Market, generate major account leads, and build client relationships*. New Jersey: John Wiley & Sons.

Gorsuch, R. (1983). Factor analysis. In J. R. C. In Nesselroade, R. B. (Eds.), (1988). Handbook of multivariate experimental psychology (2nd ed.). (231-258). NY: Plenum Press. (Ed.).

Gorsuch, R. L. (1988). Exploratory factor analysis. In J. R. Nesselroade & R. B. Cattell, (Eds.), Handbook of multivariate experimental psychology (2nd ed., pp. 231-258). NY, US: Plenum Press.

Grabe, W. (1987). Contrastive rhetoric and text-type research. *Writing across languages: Analysis of L2 text*, 115-138.

Greenbaum, S., & Quirk, R. (1990). *A student's grammar of the English language*. London: Longman.

Gregory, M. (1967). Aspects of varieties differentiation. *Journal of Linguistics, 3*(02), 177-198.

Gregory, M., & Carroll, S. (1978). *Language and situation: Language varieties and their social contexts*. London: Routledge.

Grieve, J., Biber, D., Friginal, E., & Nekrasova, T. (2011). Variation among blogs: a multi-dimensional analysis. In A. Mehler, S. Sharoff, & M. Santini(Eds.), Genres on the web: Computational models and empirical studies (pp.303-322). London: Springer.

Grimes, J. E. (1975). *The thread of discourse* (Vol. 207). Berlin: Walter de Gruyter.

Guilford, J. P. (1954). *Psychometric methods*. New York: McGraw-Hill.

Gurak, L. J., Antonijevic, S., Johnson, L., Ratliff, C., & Reyman, J. (2004). Into the blogosphere: Rhetoric, community, and culture of weblogs. *Online edited collection of scholarly articles*.

Hård af Segerstad, Y. (2002). *Use and adaptation of written language to the conditions of computer-mediated communication.* (Doctoral Dissertation), Göteborg University, Göteborg, Sweden.

Haas, T. (2005). From "Public Journalism" to the "Public's Journalism"? Rhetoric and reality in the discourse on weblogs. *Journalism Studies, 6*(3), 387-396.

Halliday, M. A. K. (1967). Notes on transitivity and theme in English (Part II). *Journal of linguistics, 3*, 199-244.

Halliday, M. A. K. (1970a). Functional diversity in language as seen from a consideration of modality and mood in English. *Foundations of language*, 322-361.

Halliday, M. A. K. (1970b). Language structure and language function. In J. Lyons (Ed.), *New horizons in linguistics* (Vol. 1, pp. 140-165). Harmondsworth, Middlesex: Penguin.

Halliday, M. A. K. (1978). Language as Social Semiotic: the social interpretation of language and meaning. *London: Arnold*.

Halliday, M. A. K. (1983). 1. Ideas About Language. *Essays in Honor of Charles F. Hockett, 4*, 15.

Halliday, M. A. K. (1985). *Spoken and written language*. Oxford: Oxford University Press.

Halliday, M. A. K. (1987). Spoken and Written Modes of Meaning. *In R. Horowitz & S. J. Samuels (Eds.), Comprehending oral and written language (pp. 55-82). New York: Academic Press*.

Halliday, M. A. K. (1988). On the language of physical science. *Registers of written English: Situational factors and linguistic features*, 162-178.

Halliday, M. A. K. (1989). *Spoken and written language*. Oxford, UK: Oxford University Press.

Halliday, M. A. K. (1992). Language as system and language as instance: the corpus as a theoretical construct. *Directions in corpus linguistics*, 61-77.

Halliday, M. A. K. (1994). Functional grammar. *London: Edward Arnold*.

Halliday, M. A. K. (1996). Literacy and linguistics: A functional perspective. *Literacy in society*, 339-376.

Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English*. London: Longman.

Halliday, M. A. K., & Hasan, R. (1989). *Language, context, and text: Aspects of language in a social-semiotic perspective*. Oxford: Oxford University Press.

Halliday, M. A. K., & Martin, J. R. (1981). *Readings in systemic linguistics*. Trafalgar Square Publishing.

Halliday, M. A. K., & Martin, J. R. (1993). *Writing science: Literacy and discursive power*. London and Washington: Palmer.

Halliday, M. A. K., & Matthiessen, C. M. I. M. (1999). Construing experience through meaning. A language-based approach to cognition. London. *Canadian Journal of Linguistics Revue Canadienne De Linguistique, 27*(1), 140-142.

Halliday, M. A. K., & Matthiessen, C. M. I. M. (2004). *An introduction to functional grammar* (Vol. 3). London: Arnold.

Halliday, M. A. K., McIntosh, A., & Strevens, P. (1964). *The linguistic sciences and language teaching*. London: Longman.

Halliday, M. A. K., & Webster, J. (2006). *On language and linguistics* (Vol. 3). London: Continuum.

Hartigan, J. A. (1985). Statistical theory in clustering. *Journal of classification, 2*(1), 63-76.

Harwood, N. (2005). 'We do not seem to have a theory… The theory I present here attempts to fill this gap': Inclusive and exclusive pronouns in academic writing. *Applied Linguistics, 26*(3), 343-375.

Hasan, R. (1979). On the notion of text. In I. J. P. (ed.) (Ed.), *Text Versus Sentence: Basic Questions of Text Linguistics. Papers in Text Linguistics 20. Hamburg: Buske. Cited in Hasan (1999).*

Hasan, R. (1984). The nursery tale as a genre. *Nottingham Linguistic Circular 13. Reprinted in Hasan (1996).*

Hasan, R. (1999). Speaking with reference to context. In M. Ghadessy (Ed.), Text and Context in Functional Linguistics: Systemic Perspectives. (pp. 219-328).

Hasan, R. (2009). The place of context in a systemic functional model. *Continuum companion to systemic functional linguistics*, 166-189.

Haspelmath, M. (1997). *From space to time: Temporal adverbials in the world's languages*. Munich: Lincom Europa.

Hatch, E. (1992). *Discourse and language education*. Cambridge: Cambridge University Press.

Hayton, J. C., Allen, D. G., & Scarpello, V. (2004). Factor retention decisions in exploratory factor analysis: A tutorial on parallel analysis. *Organizational research methods, 7*(2), 191-205.

Healthfield, S. M. (2010) Why Blogging Matters to Business. URL: https://occationalblogger.wordpress.com/2010/10/04/why-blogging-matters-to business/

Herring, S. C. (2001). Computer-mediated discourse analysis. In D. Shiffrin, D. Tannen & H.E. Hamilton (Eds.), The handbook of discourse analysis. (613-634). Malden: Blackwell Publishers Ltd.

Herring, S. C. (1996). *Computer-mediated communication: Linguistic, social, and cross-cultural perspectives* (Vol. 39). Amsterdam: John Benjamins Publishing.

Herring, S. C. (2007). A faceted classification scheme for computer-mediated discourse. *Language@ internet, 4*(1), 1-37.

Herring, S. C., & Paolillo, J. C. (2006). Gender and genre variation in weblogs. *Journal of Sociolinguistics, 10*(4), 439-459.

Herring, S. C., Scheidt, L. A., Bonus, S., & Wright, E. (2004). *Bridging the gap: A genre analysis of weblogs.* Paper presented at the Proceedings of the 37th Annual Hawaii International Conference on on System Sciences (HICSS-37). , Los Alamitos.

Hjelmslev, L. (1961). *Prolegomena to a Theory of Language*. Madison: University of Wisconsin Press.

Hockey, S. (2001). Concordance programs for corpus linguistics. In R. C. Simpson and J. M. Swales (eds.), Corpus Linguistics in North America, pp.76-97. . Ann Arbor: University of Michigan Press.

Hoey, M. (1983). *On the surface of discourse*. London: George Allen and Unwin.

Hoffmann, C. R. (2012). *Cohesive profiling: Meaning and interaction in personal weblogs* (Vol. 219): John Benjamins Publishing.

Hookway, N. (2008). Entering the blogosphere': some strategies for using blogs in social research. *Qualitative research, 8*(1), 91-113.

Horvath, B. M. (1985). Variation in Australian English. The sociolects of Sydney. *Cambridge Studies in Linguistics London*(45), 1-200.

Hudson-Ettle, D. M. (1998). *Grammatical subordination strategies in differing text types in the English spoken and written in Kenya*. Unpublished PhD thesis.

Huffaker, D. A., & Calvert, S. L. (2005). Gender, identity, and language use in teenage blogs. *Journal of Computer-Mediated Communication, 10*(2).

Hunston, S. (2002). *Corpora in applied linguistics*. Cambridge: Cambridge University Press.

Hunston, S., & Francis, G. (2000). *Pattern grammar: A corpus-driven approach to the lexical grammar of English*. Amsterdam: John Benjamins.

Hunston, S., & Thompson, G. (2000). *Evaluation in Text: Authorial Stance and the Construction of Discourse: Authorial Stance and the Construction of Discourse*. Oxford: Oxford University Press.

Hunt, K. W. (1965). Grammatical Structures Written at Three Grade Levels. NCTE Research Report No. 3.

Hutchinson, R. (2001). *Dangerous liaisons? Ethical considerations in conducting online sociological research.* Paper presented at the TASA 2001 Conference (The Australian Sociological Association 13 December 2001 to 15 December 2001).

Hyland, K. (1996). Talking to the academy forms of hedging in science research articles. *Written communication, 13*(2), 251-281.

Hyland, K. (2001). Humble servants of the discipline? Self-mention in research articles. *English for Specific purposes, 20*(3), 207-226.

Hyland, K. (2002). Authority and invisibility: Authorial identity in academic writing. *Journal of pragmatics, 34*(8), 1091-1112.

Hymes, D. (1974). *Foundations in sociolinguistics: An ethnographic approach*. Philadelphia: University of Pennsylvania Press.

Hymes, D. (1984). Sociolinguistics: stability and consolidation. *International Journal of the Sociology of Language, 1984*(45), 39-46.

Johansson, S. (2012). *Interpreting textual distribution: social and situational factors*: na.

Johansson, V. (2008). Lexical diversity and lexical density in speech and writing. *Lund University, Department of Linguistics and Phonetics*.

Johnson, R. A., & Wichern, D. W. (1992). *Applied multivariate statistical analysis* (Vol. 4). NJ: Prentice hall Englewood Cliffs.

Joos, M. (1967). *The five clocks*. New York: Harcourt, Brace & World.

Kühnen, U., Hannover, B., & Schubert, B. (2001). The semantic–procedural interface model of the self: The role of self-knowledge for context-dependent versus context-independent modes of thinking. *Journal of personality and social psychology, 80*(3), 397.

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and psychological measurement, 20*, 141-151.

Kaiser, S., & Müller-Seitz, G. (2008). Leveraging lead user knowledge in software development—The case of weblog technology. *Industry and Innovation, 15*(2), 199-221.

Kashima, E. S., & Kashima, Y. (1998). Culture and Language The Case of Cultural Dimensionsand Personal Pronoun Use. *Journal of Cross-Cultural Psychology, 29*(3), 461-486.

Kelleher, T., & Miller, B. M. (2006). Organizational blogs and the human voice: Relational strategies and relational outcomes. *Journal of Computer-Mediated Communication, 11*(2), 395-414.

Kennedy, G. (1998). *An introduction to corpus linguistics*. London: Longman.

Kern, M. C., Lee, S., Aytug, Z. G., & Brett, J. M. (2012). Bridging social distance in inter-cultural negotiations:"you" and the bi-cultural negotiator. *International Journal of Conflict Management, 23*(2), 173-191.

Kern, R. (2006). Perspectives on technology in learning and teaching languages. *Tesol Quarterly, 40*(1), 183-210.

Kim, J.-O., & Mueller, C. W. (1978). *Factor analysis: Statistical methods and practical issues*. Newbury Park, CA: Sage Publications.

King, S. A. (1996). Researching Internet communities: Proposed ethical guidelines for the reporting of results. *The Information Society, 12*(2), 119-128.

Kiousis, S. (2001). Public trust or mistrust? Perceptions of media credibility in the information age. *Mass Communication & Society, 4*(4), 381-403.

Klare, G. R. (1963). *Measurement of readability*. Ames, IA: Iowa State University Press.

Klein, D., & Manning, C. D. (2002). *Fast exact inference with a factored model for natural language parsing.* Paper presented at the Advances in neural information processing systems.

Klein, D., & Manning, C. D. (2003). *Accurate unlexicalized parsing.* Paper presented at the Proceedings of the 41st Annual Meeting on Association for Computational Linguistics-Volume 1.

Kline, P. (1979). *Psychometrics and psychology*. London, UK: Academic Press.

Koenig Kellas, J., Baxter, L., & Braithwaite, D. (2008). Narrative theories: Making sense of interpersonal communication. *Engaging theories in interpersonal communication: Multiple perspectives*, 241-254.

Krishnamurthy, S. (2002). The multidimensionality of blog conversations: The virtual enactment of September 11. 3th International and Interdisciplinary Association of Internet Research 3.0. *Netherlands, Maastricht*.

Kroch, A. S., & Hindle, D. M. (1982). Quantitative Study of the Syntax of Speech and Writing *Technical Report*. Pennsylvania: University of Pennsylvania.

Kuo, C.-H. (1999). The use of personal pronouns: Role relationships in scientific journal articles. *English*

for specific purposes, 18(2), 121-138.

Kwasnik, B. H., Crowston, K., Herring, S. C., Scheidt, L. A., Wright, E., & Bonus, S. (2005). Weblogs as a bridging genre. *Information Technology & People, 18*(2), 142-171.

Lasica, J. D. (2001). Blogging as a form of journalism. *Online Journalism Review, 24*.

Lavid, J. (1995). Towards a text type taxonomy: a functional framework for text analysis and generation. *Rev. Procesaminento Lenguaje Natural, 16*, 29-43.

Lee, D. Y. (2001). Genres, registers, text types, domains and styles: Clarifying the concepts and nevigating a path through the BNC jungle. *Language Learning & Technology, 5*(3), 37-72.

Lee, H. B., & Comrey, A. L. (1979). Distortions in a commonly used factor analytic procedure. *Multivariate Behavioral Research, 14*(3), 301-321.

Leech, G. (2009). *Change in contemporary English: a grammatical study*. Cambridge: Cambridge University Press.

Leech, G. N. (1971). *Meaning and the English verb*. Harlow, England: Longman.

Leech, G. N. (1991). The state of the art in corpus linguistics. In K. In Aijmer, and B. Altenberg (Ed.) English Corpus Linguistics: Studies in Honor of Jan Svartvik (Ed.), (pp. 8-29). London: Longman.

Lenhart, A., & Fox, S. (2006). Bloggers. Pew Internet & American Life Project. *Retrieved March, 8*, 2011.

Levy, R., & Andrew, G. (2006). *Tregex and Tsurgeon: tools for querying and manipulating tree data structures.* Paper presented at the Proceedings of the fifth international conference on Language Resources and Evaluation.

Li, L.-J., & Ge, G.-C. (2009). Genre analysis: Structural and linguistic evolution of the English-medium medical research article (1985–2004). *English for Specific Purposes, 28*(2), 93-104.

Lixia, M. (2004). The Application of Blog in Organizational Communication. *Information Science, 7*.

Lu, X. (2010). Automatic analysis of syntactic complexity in second language writing. *International Journal of Corpus Linguistics, 15*(4), 474-496.

Ludeling, A., & Kyto, M. (2008). *Corpus linguistics: An international handbook*. Berlin: Mouton de.

Lukin, A., Moore, A., Herke, M., Wegener, R., & Wu, C. (2008). Halliday's model of register revisited and explored. *Linguistics & the Human Sciences*(*4*), 187-213.

Luppicini, R. (2007). Review of computer mediated communication research for education. *Instructional science, 35*(2), 141-185.

Lyons, J. (1977). *Semantics*. Cambridge, England: Cambridge University Press.

Mühlhäusler, P., & Harré, R. (1990). *Pronouns and people: The linguistic construction of social and personal identity*: Blackwell.

MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological methods, 4*(1), 84.

Malinowski, B. (1923). The problem of meaning in primitive languages. In C. K. O. a. I. A. Richards (Ed.), *The Meaning of Meaning* (pp. 296-336). New York: Harcourt Brace & World.

Marshall, I. (1987). Tag selection using probabilistic methods. *The Computational analysis of English: a corpusbased approach*, 42-65.

Martin, J. R. (1997). Analyzing genre: functional parameters. In F. Christie, F. & J. R. Martin (eds) Genre and Institutions - Social Processes in the Workplace and School. (pp. 3-39). London; New York: Continuum.

Martin, J. R. (1984). Language, register and genre. *Children Writing: Reader, 1*, 984.

Martin, J. R. (1985). Process and text: two aspects of human semiosis. *Systemic perspectives on*

discourse, 1(15), 248-274.

Martin, J. R. (1991). Intrinsic functionality: implications for contextual theory. *Social semiotics, 1*(1), 99-162.

Martin, J. R. (1992). *English text: System and structure*. Amsterdam: John Benjamins Publishing.

Martin, J. R. (1999a). Mentoring semogenesis:'genre-based'literacy pedagogy. *Pedagogy and the shaping of consciousness: Linguistic and social processes*, 123-155.

Martin, J. R. (1999b). Modelling context. *Text and context in functional linguistics*, 25-62.

Martin, J. R. (2010). Semantic variation: modelling system, text and affiliation in social semiosis. In M. B. J. R. Martin (Ed.), *New discourse on language: Functional perspectives on multimodality, identity and affiliation* (pp. 1-34). London: Continuum.

Martin, J. R., & Rose, D. (2003). *Working with discourse: Meaning beyond the clause*. London: Bloomsbury Publishing.

Martin, J. R., & Rose, D. (2007). Interacting with text: The role of dialogue in learning to read and write. *Foreign Studies Journal*.

Martin, J. R., & Rose, D. (2008). *Genre relations: Mapping culture*. London: Equinox.

Maslansky, M., West, S., DeMoss, G., & Saylor, D. (2010). *The language of trust: selling ideas in a world of skeptics*. New York: Prentice Hall Press.

Matthiessen, C. M. (2007). The "architecture" of language according to systemic functional theory: developments since the 1970s *In R. Hasan, C. M. I. M. Matthiessen, & J. J. Webster (Eds.), Continuing discourse on language* (Vol. 2, pp. 505-561). London: Equinox Publishing.

Matthiessen, C. M., Teruya, K., & Canzhong, W. (2008). Multilingual studies as a multi-dimensional space of interconnected language studies. *Meaning in context: implementing intelligent applications of language studies*, 146-220.

McEnery, T., & Wilson, A. (2001). *Corpus linguistics: An introduction*. Edinburgh: Edinburgh University Press.

McEnery, T., Xiao, R., & Tono, Y. (2006). *Corpus-based language studies: An advanced resource book*. London: Routledge.

McGregor, J. (2004). It's a blog world after all. *Fast Company*, 84-87.

Miller, C. R., & Shepherd, D. (2004). Blogging as social action: A genre analysis of the weblog. *Into the blogosphere: Rhetoric, community, and culture of weblogs, 18*(1), 1-24.

Milligan, G. W., & Cooper, M. C. (1985). An examination of procedures for determining the number of clusters in a data set. *Psychometrika, 50*(2), 159-179.

Milroy, J., & Milroy, L. (2005). Language in society: Sociolinguistics. *An encyclopaedia of language*, 267.

Moessner, L. (2001). Genre, text type, style, register: A terminological maze? *European Journal of English Studies, 5*(2), 131-138.

Murray, D. E. (1988). The Context of Oral and Written Language: A Framework for Mode and Medium Switching. *Language in Society, 17*(3), 351-373.

Murray, D. E. (1990). CmC. *English Today, 6*(03), 42-46.

Myers, G. (2010). *The discourse of blogs and wikis*. London: Continuum.

Nakajima, S., Tatemura, J., Hara, Y., Tanaka, K., & Uemura, S. (2006). Identifying agitators as important blogger based on analyzing blog threads *In APWeb 2006* (pp. 285-296).

Nelson, G. (1996). The design of the corpus. *Comparing English Worldwide: The International Corpus of English*, 27-35.

Nelson, M. (2006). Semantic associations in Business English: A corpus-based analysis. *English for*

Specific Purposes, 25(2), 217-234.

Nelson, M. B., & West, R. (2000). *corpus-based study of the lexis of business English and business English teaching materials.* (Unpublished PhD thesis), University of Manchester.

Nesi, H. (2011). BAWE: an introduction to a new resource. *New trends in corpora and language learning*, 213-228.

Nesi, H., & Gardner, S. (2012). *Genres across the disciplines: Student writing in higher education*: Cambridge University Press.

Neumann, S. (2010). Quantitative register analysis across languages. In Swain, Elizabeth (ed.), Thresholds and Potentialities of Systemic Functional Linguistics: Applications to other disciplines, specialised discourses and languages other than English. Trieste: Edizioni Universitarie.

Nguyen, L. V. (2008). Computer mediated communication and foreign language education: Pedagogical features. *International journal of instructional technology and distance learning, 5*(12), 23-44.

Nilsson, S. (2003). The function of language to facilitate and maintain social networks in research weblogs. *D-Essay. Umea Universitet, Engelska lingvistik. Retrieved February, 27*, 2004.

Nini, A., & Grant, T. (2013). Bridging the gap between stylistic and cognitive approaches to authorship analysis using Systemic Functional Linguistics and multidimensional analysis. *International Journal of Speech, Language & the Law, 20*(2).

Nowson, S., Oberlander, J., & Gill, A. J. (2005). *Weblogs, genres and individual differences.* Paper presented at the Proceedings of the 27th Annual Conference of the Cognitive Science Society.

Nunnally, J. C., & Bernstein, I. H. (1978). Psychometric theory. New York: McGraw-Hill.

Nwogu, K. N. (1989). *Discourse variation in medical texts: Schema, theme and cohesion in professional and journalistic accounts.* Aston University.

O'Donnell, R. C. (1974). Syntactic differences between speech and writing. *American speech*, 102-110.

Ochs, E., & Capps, L. (2009). *Living narrative: Creating lives in everyday storytelling*. Cambridge: Harvard University Press.

Olekalns, M., & Smith, P. L. (2003). Testing the relationships among negotiators' motivational orientations, strategy choices, and outcomes. *Journal of Experimental Social Psychology, 39*(2), 101-117.

Ortega, L. (2003). Syntactic complexity measures and their relationship to L2 proficiency: A research synthesis of college‐level L2 writing. *Applied Linguistics, 24*(4), 492-518.

Palmer, F. R. (2014). *The english verb*. New York: Routledge.

Paltridge, B. (1996). Genre, text type, and the language learning classroom. *ELT journal, 50*(3), 237-243.

Park, D. W. (2009). Blogging with authority: Strategic positioning in political blogs. *International Journal of Communication, 3*, 24.

Pennycook, A. (1994). The politics of pronouns. *ELT journal, 48*(2), 173-178.

Phelan, J., & Rabinowitz, P. J. (2008). *A companion to narrative theory*. Oxford: Blackwell.

Pike, K. L. (1954). *Language in relation to a unified theory of the structure of human behavior*. Glendale, CA: Summer Institute of Linguistics.

Potts, K. (2007). *Web Design and Marketing Solutions for Business Websites*. New York: Apress.

Poynton, C. (1989). *Language and gender: Making the difference*. Oxford: Oxford University Press.

Puschmann, C. (2010a). *The corporate blog as an emerging genre of computer-mediated communication: features, constraints, discourse situation* (Vol. 7): Universitätsverlag Göttingen.

Puschmann, C. (2010b). Thank you for thinking we could. *Use and function of interpersonal pronouns in corporate web logs. In Approaches to Syntactic Variation and Genre, Topics in English*

Linguistics. Mouton de Gruyter.

Quible, Z. K. (2005). Blogs and written business communication courses: A perfect union. *Journal of Education for Business, 80*(6), 327-332.

Quirk, R., Crystal, D., & Education, P. (1985). *A comprehensive grammar of the English language* (Vol. 397). Cambridge: Cambridge University Press.

Quirk, R., & Greenbaum, S. (1973). *A university grammar of English*. London: Longman.

Quirk, R., & Greenbaum, S. (1993). *A University Grammar of English*. London: Longman Group UK Ltd.

Ravid, D., & Berman, R. A. (2006). Information density in the development of spoken and written narratives in English and Hebrew. *Discourse Processes, 41*(2), 117-149.

Reise, S. P., Waller, N. G., & Comrey, A. L. (2000). Factor analysis and scale revision. *Psychological assessment, 12*(3), 287.

Reppen, R. (2001). Register variation in student and adult speech and writing. *In S. Conrad and D. Biber (eds) Variation in English: Multi–dimensional studies. London: Longman, 187–199.*

Resnick, S. A. M. P. (2011). *The prevalence of political discourse in non-political blogs*. London: Routledge. .

Rev, I. (2008). *4.0. International Standard Industrial Classification of All Economic Activities. Revision 4*. United Nations, New York.

Rimmon-Kenan, S. (1983). *Narrative fiction: Contemporary poetics*. London: Routledge.

Robinson, S. (2006). The mission of the j-blog Recapturing journalistic authority online. *Journalism, 7*(1), 65-83.

Rounds, P. L. (1985). *Talking the mathematics through: Disciplinary transaction and socio-educational interaction.* University of Michigan.

Rowley-Jolivet, E., & Carter-Thomas, S. (2005). Genre awareness and rhetorical appropriacy: Manipulation of information structure by NS and NNS scientists in the international conference setting. *English for Specific Purposes, 24*(1), 41-64.

Rutherford, W. E. (2014). *Second language grammar: Learning and teaching*. London: Routledge.

Santini, M. (2005). Linguistic facets for genre and text type identification: A description of linguistically-motivated features. *ITRI report series: ITRI-05, 2*.

Santini, M. (2006). Web pages, text types, and linguistic features: Some issues. *Icame Journal, 30*, 67-86.

SARDINHA, T. B., KAUFFMAN, C., & ACUNZO, C. M. Dimensions of variation in spoken Brazilian Portuguese.

Scale, C. (2002). 2 Resurrective practice and narrative. *Lines of narrative: Psychosocial perspectives*, 36.

Schleppegrell, M. J. (2002). Linguistic features of the language of schooling. *Linguistics and education, 12*(4), 431-459.

Schleppegrell, M. J. (2004). *The language of schooling: A functional linguistics perspective*. Mahwah, NJ: Erlbaum.

Schmidt, J. (2007). Blogging practices: An analytical framework. *Journal of Computer ‐Mediated Communication, 12*(4), 1409-1427.

Sharda, N., & Ponnada, M. (2008). Tourism blog visualizer for better tour planning. *Journal of Vacation Marketing, 14*(2), 157-167.

Sharoff, S. (2006). Creating general-purpose corpora using automated search engine queries. *WaCky*, 63-98.

Sifry, D. (2004). State of the Blogosphere. www.sifry.com/alerts/archives/2004_10.html. .

Simmons, R. A., Gordon, P. C., & Chambless, D. L. (2005). Pronouns in Marital Interaction What Do "You" and "I" Say About Marital Health? *Psychological science, 16*(12), 932-936.

Simpson, P. (1993). *Language, Ideology, and Point of View*. London: Routledge.

Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford: Oxford University Press.

Sinclair, J. (2005). Corpus and text-basic principles. *Developing linguistic corpora: A guide to good practice*, 1-16.

Sinclair, J. M., & Coulthard, M. (1975). *Towards an analysis of discourse: The English used by teachers and pupils*. Oxford: Oxford Univ Press.

Singer, J. B. (2005). The political j-blogger 'normalizing'a new media form to fit old norms and practices. *Journalism, 6*(2), 173-198.

Singh, R. P., & Singh, L. O. (2008). Blogs: Emerging knowledge management tools for entrepreneurs to enhance marketing efforts. *Journal of Internet Commerce, 7*(4), 470-484.

Sixsmith, J., & Murray, C. D. (2001). Ethical issues in the documentary data analysis of Internet posts and archives. *Qualitative Health Research, 11*(3), 423-432.

Smith, E. L. (1982). Functional types of scientific prose. In J. D. B. W. S. Greaves (Ed.), *Systemic perspectives on discourse* (pp. 241-257). Norwood, NJ: Ablex.

Spiegelberg, H. (1973). On the right to say 'we': A linguistic and phenomenological analysis. *Phenomenological sociology*, 129-156.

Steiner, E. (2004). *Translated texts: properties, variants, evaluations*. Frankfurt: Peter Lang Frankfurt.

Stuart, K. (2006). Towards an analysis of academic weblogs. *Revista alicantina de estudios ingleses, No. 19 (Nov. 2006); pp. 387-404*.

Stubbs, M. (1996). *Text and corpus analysis: Computer-assisted studies of language and culture*. Oxford: Blackwell

Sturgess, P. J. M. (1992). *Narrativity: theory and practice*. Oxford: Clarendon Press.

Sudweeks, F., & Rafaeli, S. (1996). How Do You Get a Hundred Strangers to Agree? *Computer Networking and Scholarly Communication in the Twenty-First Century, TM Harrison and T. Stephen, eds., Albany: State University of New York*, 115-136.

Suitt, H. (2003). A blogger in their midst. *Harvard business review, 81*(9), 30-33; discussion 34-36, 38, 40 passim.

Swales, J. (1990). *Genre analysis: English in academic and research settings*. Cambridge: Cambridge University Press.

Tagliacozzo, R. (1976). Levels of technicality in scientific communication. *Information Processing & Management, 12*(2), 95-110.

Tannen, D. (1982a). Oral and literate strategies in spoken and written narratives. *Language*, 1-21.

Tannen, D. (1982b). *Spoken and written language: Exploring orality and literacy* (Vol. 32): Ablex Publishing Corporation.

Tannen, D. (1985). *Relative focus on involvement in oral and written discourse.* In D. R. Olson, N. Torrance, & A. Hildyard (Eds.), Literacy, language, and learning. Cambridge: Cambridge University Press.

Tavosanis, M. (2006). *Linguistic Features of Italian Blogs.* Paper presented at the the Proceedings of the Workshop on NEW TEXT: Wikis and Blogs and Other Dynamic Text Sources, EACL.

Tavosanis, M. (2007). *Juvenile Netspeak and subgenre classification issues in Italian blogs.* Paper presented at the Proc. of the Int. Workshop Towards Genre-Enabled Search Engines.

Teich, E. (2003). *Cross-linguistic variation in system and text: A methodology for the investigation of translations and comparable texts* (Vol. 5). Berlin: Walter de Gruyter.

Thompson, G. (1996). *Introduction to functional grammar*. London: Edward Arnold.

Thompson, G. (2004). *Introducing functional grammar*. London: Arnold.

Thompson, G. (2013). *Introducing functional grammar*. London: Routledge.

Thompson, S. A. (1983). Grammar and discourse: The English detached participial clause. *Discourse perspectives on syntax*, 43-65.

Thompson, S. A. (2002). "Object complements" and conversation: Towards a realistic account. *Studies in language, 26*(1), 125-163.

Thorne, S. L. (2008). Mediating technologies and second language learning. *Handbook of research on new literacies*, 417-449.

Tirapat, T., Espiritu, C., & Stroulia, E. (2006). *Taking the community's pulse: one blog at a time.* Paper presented at the Proceedings of the 6th international conference on Web engineering.

Torres-Salinas, D., Cabezas-Clavijo, Á., Ruiz-Pérez, R., & López-Cózar, E. D. (2011). State of the library and information science blogosphere after social networks boom: A metric approach. *Library & information science research, 33*(2), 168-174.

Tsfati, Y. (2003). Does audience skepticism of the media matter in agenda setting? *Journal of Broadcasting & Electronic Media, 47*(2), 157-176.

Tsui, A. B. (1991). Sequencing rules and coherence in discourse. *Journal of pragmatics, 15*(2), 111-129.

Tsui, A. B. (1994). *English conversation*: Oxford Univ Pr.

Tucker, L. R., Koopman, R. F., & Linn, R. L. (1969). Evaluation of factor analytic research procedures by means of simulated correlation matrices. *Psychometrika, 34*(4), 421-459.

Ure, J. (1969). Practical registers. *ELT journal, 23*(3), 206-215.

Ure, J. (1971). Lexical density and register differentiation. *Applications of linguistics*, 443-452.

Ure, J., & Ellis, J. (1977). Register in descriptive linguistics and linguistic sociology. *Issues in Sociolinguistics. The Hague: Mouton*, 197-243.

Van Dijk, T. A. (2008). *Discourse and context: A Socio-cognitive Approach*. Cambridge: Cambridge University Press.

Van Rooy, B., & Terblanche, L. (2009). A multi-dimensional analysis of a learner corpus. *Language and Computers, 69*(1), 239-254.

Velicer, W. F., Eaton, C. A., & Fava, J. L. (2000). Construct explication through factor or component analysis: A review and evaluation of alternative procedures for determining the number of factors or components. In R. D. Goffin & E. Helms (Eds.), Problems and solutions in human assessment: Honoring Douglas N. Jackson at seventy (pp. 41-71). Norwell, MA: Springer.

Velicer, W. F., & Fava, J. L. (1987). An evaluation of the effects of variable sampling on component, image, and factor analysis. *Multivariate Behavioral Research, 22*(2), 193-209.

Velicer, W. F., & Fava, J. L. (1998). Affects of variable and subject sampling on factor pattern recovery. *Psychological methods, 3*(2), 231.

Ventola, E. (1984). Orientation to Social Semiotics in Foreign Language Teaching1. *Applied Linguistics, 5*(3), 275-286.

Verstraete, J.-C. (2001). Subjective and objective modality: Interpersonal and ideational functions in the English modal auxiliary system. *Journal of pragmatics, 33*(10), 1505-1528.

Virtanen, T. (1990). On the definitions of text and discourse. *Folia Linguistica, 24*(3-4), 447-455.

Virtanen, T. (1992). Issues of text typology: Narrative—a 'basic'type of text? *Text-Interdisciplinary Journal*

for the Study of Discourse, 12(2), 293-310.

Vobič, I. (2007). The normalization of the blog in journalism: online newspapers of Slovene traditional media. *Medijska istraživanja, 13*(2), 59-82.

Vold, E. T. (2006). Epistemic modality markers in research articles: a cross-linguistic and cross-disciplinary study. *International Journal of Applied Linguistics, 16*(1), 61-87.

Wallsten, K. (2007). Agenda setting and the blogosphere: An analysis of the relationship between mainstream media and political blogs. *Review of Policy Research, 24*(6), 567-587.

Wardhaugh, R. (1985). *How conversation works*. Oxford: Basil Blackwell.

Waskul, D. (1996). Considering the electronic participant: Some polemical observations on the ethics of on-line research. *The Information Society, 12*(2), 129-140.

Weil, D. (2006). *The corporate blogging book*. London: Piatkus.

Weingart, L. R., Thompson, L. L., Bazerman, M. H., & Carroll, J. S. (1990). Tactical behavior and negotiation outcomes. *International Journal of Conflict Management, 1*(1), 7-31.

Werlich, E. (1976). *A text grammar of English*. Heidelberg: Quelle & Meyer.

Werlich, E. (1983). *A Text Grammar of English*. Heidelberg: Quelle & Meyer.

Werry, C. C. (1996). Linguistic and Interactional Features of Internet Relay Chat. I Herring, S.(red.): Computer-Mediated Communication: Linguistic, Social and Cross-Cultural Perspectives (pp. 47–63). *Amsterdam: John Benjamin*.

Westin, I. (2002a). Features marking information density. *Language and Computers, 44*(1), 65-84.

Westin, I. (2002b). *Language change in English newspaper editorials.* Amsterdam: Rodopi.

Westin, I., & Geisler, C. (2002). A multi-dimensional study of diachronic variation in British newspaper editorials. *ICAME journal, 26*, 133-152.

White, M. (1994). *Language in job interviews: Differences relating to success and socioeconomic variables.* (PhD dissertation), Northern Arizona University.

White, M. (2003). Web Logs: Moving Beyond Cool. *Econtent, 26*(12), 10-10.

Widdowson, H. G. (1983). *Learning purpose and language use*. Oxford: Oxford University Press.

Williams, J. B., & Jacobs, J. (2004). Exploring the use of blogs as learning spaces in the higher education sector. *Australasian journal of educational technology, 20*(2).

Winter, E. O. (1982). *Towards a contextual grammar of English: the clause and its place in the definition of sentence*. London: Allen & Unwin.

Wolfe-Quintero, K., Inagaki, S., & Kim, H.-Y. (1998). *Second language development in writing: Measures of fluency, accuracy, & complexity*: University of Hawaii Press.

Wood, J. M., Tataryn, D. J., & Gorsuch, R. L. (1996). Effects of under-and overextraction on principal axis factor analysis with varimax rotation. *Psychological methods, 1*(4), 354.

Xiao, R. (2009). Multidimensional analysis and the study of world Englishes. *World Englishes, 28*(4), 421-450.

Yang, S.-U., Kang, M., & Johnson, P. (2010). Effects of narratives, openness to dialogic communication, and credibility on engagement in crisis communication through organizational blogs. *Communication Research*.

Yates, J., & Orlikowski, W. J. (1993). *Knee-jerk anti-loopism and other e-mail phenomena: Oral, written, and electronic patterns in computer-mediated communication*: Alfred P. Sloan School of Management, Massachusetts Institute of Technology.

Yates, S. J. (1996). Oral and written linguistic aspects of computer conferencing. *Pragmatics and beyond New Series*, 29-46.

Yoo, S. W., & de Zúñiga, H. G. (2014). Connecting blog, Twitter and Facebook use with gaps in knowledge and participation. *Comunicación y Sociedad, 27*(4), 33.

Young, D. (1985). Some applications of systemic grammar to TEFL, or whatever became of register analysis. *Systemic perspectives on discourse, 2*, 282-294.

Zipf, G. K. (1949). Human Behavior and the Principle of Least Effort. *Journal of Clinical Psychology, 6*(3), 306-306.

Appendix 1: Titles of texts in the CBC

No.	Titles	Launch time	Author(s)
1	Say Hello To Mobile Baggage Tracking	2011/11/22	Jonathan Mayfield
2	New Boarding Pass Debuts On Delta.com & Airport Kiosks	2012/2/1	Rick E.
3	Wi-Fi Update – The Internet Flies Delta Connection!	2011/9/7	Chris B.
4	Groupon A Delta Sky Club Pass & Go!	2011/10/25	Joy H.
5	More Economy Comfort Means More Room to Relax!	2011/10/20	Chris B.
6	The New Delta BlackBerry App is Here!	2011/10/4	Jonathan Mayfield
7	Delta's 767-300ER Fleet Getting a New Look	2011/8/5	Chris B.
8	View from 30,000 Feet: Delta on Diversity	2011/6/23	Trebor Banstetter
9	Military Travel: Baggage Policies & Our Thoughts	2011/6/7	Rachael R
10	Economy Comfort Installation Updates	2011/5/5	Chris B
11	Relax, Stretch Out & Get Comfortable	2011/2/7	Chris B
12	How To: Rebook Your Travel Plans	2011/1/7	Jerry F
13	Boeing 747 Enhancements: Flat Beds, Touch Screen TVs and More	2010/9/2	Mike B
14	We're Listening. And Delivering.	2010/7/29	Jeff Robertson
15	25 Years of MaDDogs	2012/4/30	Laura M.
16	Your Trip (Extras) – Your Choice	2012/3/13	Cheryl B
17	Behind-the-Wing: Passenger Stories	2012/4/19	Laura M.
18	Behind-the-wings: The Pilot Life At Delta	2012/3/23	Dawn C
19	A New Year Means More Ways Than Ever to Surf the Internet on Delta!	2012/1/6	Chris B.
20	Fare Displays On delta.com	2012/5/18	Bob Kupbens
21	Cooking With Almond Butter	2012/2/20	Alana Sugar
22	Just Beet It	2012/2/16	Mary Olivar
23	Help a Million More People Change Their Lives	2012/2/22	Whole Planet Foundation
24	Eating Near the Poverty Line…at Whole Foods Market	2012/2/22	Lisa Johnson
25	Sweet Recipes for Sweet Potatoes	2012/2/13	Alana Sugar
26	Putumayo Celebrates Carnival in Rio with Brazilian Beat	2012/2/11	Dan Storper
27	What's New in the Grocery Aisles: February	2012/2/8	Elizabeth Smith
28	Natural Ingredients for Energy	2010/3/16	Chris Jensen
29	GMO/Monsanto Buyout Rumors Untrue	2012/2/2	Libba Letton
30	Your Favorite Holiday Food Tradition		Elizabeth Beal

		2011/12/12	
31	Win Satsumas for Sharing	2011/12/5	Jennifer Cheng
32	Enjoy A Celtic Christmas from Putumayo	2011/12/3	Dan Storper
33	The Thanksgiving I'm Most Thankful For Contest	2011/11/14	Anna Madrona
34	Seasonal Wellness Shopping List	2011/10/29	Chris Jensen
35	Let Us Stock Your Thanksgiving Pantry	201110/28	Allison Burch
36	Homeopathy for Cold and Flu Season	2011/10/18	Chris Jensen
37	Putumayo's Acoustic Café	2011/10/8	Dan Storper
38	$5 Million in Loans to Local Producers	2011/10/6	Heather Kennedy
39	Organic Valley	2011/9/20	Marc Hamel and Ha Lam
40	The Annual "Best Ever" Guacamole Search	2012/1/30	James Parker
41	AWARDS OF EXCELLENCE	2012/5/9	Bill Marriott
42	What's an Executive Chairman	2012/4/30	Bill Marriott
43	This Isn't Your Dad's Courtyard Anymore	2012/4/23	Bill Marriott
44	Cleaning up the Red Sea One Dive at a Time	2012/4/22	Bill Marriott
45	I'm Blown Away, Thanks to our GM's	2012/4/18	Bill Marriott
46	Just Plain Happiness	2012/4/9	Bill Marriott
47	The Big 8-0	2012/3/29	Bill Marriott
48	What I Learned from Starbucks' Howard Schultz	2012/3/23	Bill Marriott
49	Welcoming the World	2012/3/19	Bill Marriott
50	Building Bridges to Opportunity	2012/3/13	Bill Marriott
51	Living the American Dream	2012/3/5	Bill Marriott
52	World of Opportunity: Chance Encounters	2012/2/27	Bill Marriott
53	More Lessons From A Car Guy	2012/2/20	Bill Marriott
54	Remembering Dr. Burke, My Life-Saver	2012/2/13	Bill Marriott
55	The Super Bowl and Super Brands	2012/2/6	Bill Marriott
56	Steve Jobs: Relentless Pursuit of Excellence	2012/1/30	Bill Marriott
57	The Year of the Dragon	2012/1/23	Bill Marriott
58	A Big Day for U.S. Tourism	2012/1/20	Bill Marriott
59	Lessons of a Soda Jerk	2012/1/17	Bill Marriott
60	Monumental Courage and Conviction of Dr. Martin Luther King, Jr	2012/1/13	Bill Marriott
61	Well-Worn Wading Boots on Christmas Island	2012/5/11	Tom Morehouse
62	Ryan Montbleau Raises His Voice for Louisiana Wetlands	2012/5/9	Ryan Montbleau
63	Dirtbag Diaries: Live From 5Point Vol. 3	2012/5/8	Fitz & Becca Cahall
64	What Inspired You?	2012/5/7	
65	An Outing with Donini: Birdwatching (Part Four)	2012/5/4	Kelly Cordes
66	Gerry Lopez Australian Tour Webisode 2	2012/5/3	Patagonia Australia
67	Inside/Outside: Questions for Patagonia's Chip	2012/5/2	

	“Chipper Bro” Bell		
68	Gerry Lopez Australia Tour Webisode 1	2012/4/30	Patagonia Australia
69	An Outing with Donini: Puzzles (Part Three)	2012/4/27	Kelly Cordes
70	Patagonia Surf Stores - The Wave Riding Collective	2012/4/26	Devon Howard
71	An Interview with Photographer Ben Moon	2012/4/26	Emily Nuchols
72	Introducing the New Footprint Chronicles on Patagonia.com	2012/4/25	Lisa Polley
73	Putting Water Back	2012/4/23	Matt Stoecker & Eric Unmacht
74	Discussions of Influence	2012/4/20	Kelly Cordes
75	Heroes - Part Two	2012/4/18	Brittany Griffith
76	Justin Clifton’s 5Point of View - Catch the Film Festival April 26-29 in Carbondale	2012/4/17	James Edward Mills
77	Coral Refuge, Ocean Deep - An Excerpt from "Crossings"	2012/4/16	Michael Kew
78	An Outing with Donini: Porch Angles (Part Two)	2012/4/13	Kelly Cordes
79	Let’s Bring Back Repair	2012/4/12	Annie Leonard
80	The Wolverine Way - Go Like Hell and Never Back Down	2012/4/11	Douglas H. Chadwick
81	Toyota Sales Resurgence Gets Help from Popular Hybrid Products	2012/5/9	null
82	Toyota Motor Corporation Announced Financial Results for Fiscal Year Ended March 31, 2012	2012/5/9	null
83	Toyota to Increase 4-Cylinder Engine Capacity in Kentucky	2012/5/8	null
84	Congressional Black Caucus Foundation Names Jim Colon New Board Member	2012/5/7	null
85	All-New Toyota RAV4 EV Unveiled at EVS26	2012/5/7	null
86	2012 Electric Vehicle Symposium (EVS26) - Toyota RAV4 EV Reveal	2012/5/7	Bob Carter
87	Toyota and Discovery Education Celebrate Teachers and Students for their Creativity in Promoting Teen Driver Safety	2012/5/4	null
88	Wyland Foundation and Toyota Announce Winning Cities of Mayor's Water Conservation Challenge	2012/5/3	null
89	April 2012 Sales Conference Call	2012/5/1	Steve Curtis，Bob Carter，Tim Morrison
90	Toyota Sales Increase 11.6 Percent in April 2012	2012/5/1	null
91	Scion Announces Collaboration with its iQ and	2012/5/1	null

	Sony’s PlayStation®Vita		
92	Toyota Motor Sales Issues April 2012 Sales Preview	2012/5/1	null
93	Toyota to Debut All New RAV4 EV at EVS 26	2012/4/30	null
94	Toyota Financial Services and Lexus Financial Services Recognize Top Manheim and ADESA Auction Sites with Premier Auction Awards	2012/4/25	null
95	Toyota Named One of DiversityInc's Top 50 Companies for Diversity	2012/4/25	null
96	Toyota Announces 100 Cars for Good Finalists	2012/4/25	null
97	2013 Scion FR-S Brings the Sport Back to the Car	2012/4/25	null
98	Toyota and Wyland Foundation Recognize 30 Environmental Heroes in Southern California	2012/4/24	null
99	Toyota Announces Pricing for Limited Edition Tacoma Pickup TRD T/X Baja Series Package	2012/4/23	null
100	Toyota Racing Introduces Scion FR-S Speedster	2012/4/12	null
101	Happy National Nurses Week!	2012/5/10	Erin Ferris
102	How to Save a Life	2012/5/9	Erin Ferris
103	Weekly Worldwide Wrap-Up	2012/5/9	Nate Warren
104	Henry Dunant’s Fountain of Youth	2012/5/8	Kristiana Almeida
105	Weekly Worldwide Wrap-Up	2012/5/3	Nate Warren
106	Get Inspired to Prepare (for the Zombie Apocalypse)	2012/5/2	Gloria Huang
107	Some Like it Hot: I Like Ice	2012/5/1	Kristiana Almeida
108	Cactus Jack	2012/4/27	Wendy Harman
109	I Feel the Temperature Rising	2012/4/26	Erin Ferris
110	Weekly Worldwide Wrap-Up	2012/4/25	Nate Warren
111	Beating “The Last Measle”	2012/4/24	Kristiana Almeida
112	Serving Military Members, Veterans, and Their Families	2012/4/20	Sherri Brown
113	Weekly Worldwide Wrap-Up	2012/4/19	Nate Warren
114	Be Still My Crossword Heart	2012/4/17	Gail J McGovern
115	My Time at TEDMED	2012/4/16	Gail J McGovern
116	National Volunteer Week 2012	2012/4/16	Erin Ferris
117	Did Tornado Alley Grow?	2012/4/13	Kristiana Almeida
118	Weekly Worldwide Wrap-Up	2012/4/11	Nate Warren
119	What the Titanic Taught Me: A Lesson in Hypothermia	2012/4/11	Erin Ferris
120	The Blankets Vs. Pillows War	2012/4/10	Kristiana Almeida
121	Carousel Cakes Talks About the Busy Time of Mother’s Day	2012/5/10	Nancy Finkelman
122	Business Expertise-Customer Analytics	2012/5/9	Sally Davenport
123	EarthSmart Growing Strong within Our	2012/5/9	Philip Blackett

	Communities		
124	Summer Driving Safety	2012/5/7	Cynthia Chandler-Snell
125	EarthSmart Outreach	2012/5/4	Mitch Jackson
126	More Access Equals More Opportunity For Women	2012/5/3	Tracee Smith
127	Finding Meaning in Simple Acts	2012/4/30	Lynne Henwood
128	Daydream in Green!	2012/4/27	Lisa Watkins
129	FedEx Trade Networks Celebrates its First Year of Operations in Turkey	2012/4/26	Fred Schardt
130	FedEx Volunteers Help Beautify Shelby Farms Park in Memphis	2012/4/23	Rachel Kesselman
131	FedEx Express Establishes Default Temperature Settings on the B777F	2012/4/19	Richard Smith
132	EarthSmart @ Work	2012/4/17	Mitch Jackson
133	Cold Chain for Beginners	2012/4/17	Christelle Laot
134	Small business virtual roundtable: Disaster Preparedness - live from Joplin, Missouri on May 22	2012/4/16	Shane O'Connor
135	Reach Out and PLANT…Somebody's TREE for EarthSmart	2012/4/13	Philip Blackett
136	The Heart of America and FedEx Present School with Bright, New Reading Corner	2012/4/11	Null
137	FedEx Carbon-Neutral Envelope Shipping	2012/4/10	Mitch Jackson
138	Your Small-Business Site for LTL Shipping Success: Freight Central	2012/4/3	Logan Andres
139	FedEx team members and the 2012 Women In Aviation International Conference	2012/4/2	Null
140	How to Renew America's Economy	2012/3/29	Fred Smith
141	Introducing New Apps for Timeline	2012/1/18	Carl Sjogreen
142	Listen to Music With Your Friends	2012/1/12	Alexandre Roche
143	Timeline: Now Available on Mobile	2011/12/15	Mick Johnson
144	Timeline: Now Available Worldwide	2011/12/15	Paul McDonald
145	A Faster Facebook for Android	2011/12/7	Keith Peiris
146	Like My Status: Memology 2011	2011/12/6	Jonathan Chang
147	Our Commitment to the Facebook Community	2011/11/29	Mark Zuckerberg
148	Bigger, Faster Photos	2011/11/16	Justin Shaffer
149	Interesting News, Any Time You Visit	2011/11/9	Mark Tonkelowitz
150	A Faster Way to Message on Mobile	2011/10/19	Lucy Zhang
151	Introducing Facebook for iPad	2011/10/10	Leon Dubinsky
152	Tell Your Story with Timeline	2011/9/22	Samuel W. Lessin
153	Introducing the Subscribe Button	2011/9/14	Zach Rait
154	Improved Friend Lists	2011/9/13	Blake Ross

155	Making It Easier to Share With Who You Want	2011/8/23	Chris Cox
156	Making Games Better	2011/8/11	Jared Morgenstern
157	Call Your Friends Right From Facebook	2011/7/6	Philip Su
158	Making Photo Tagging Easier	2011/6/30	Justin Mitchell
159	Deals: Better with Friends	2011/4/26	Emily C White
160	Sharing with Small Groups	2011/4/25	Elliot Lynde
161	A little Valentine's Day inspiration (in case you need it)	2012/2/14	Turkey Hill Team
162	Which ice cream record should we attempt?	2012/1/20	Turkey Hill Team
163	The Ultimate Flavor Tournament is back! (Pick the final four flavors to win a year's supply of ice cream!)	2012/2/27	Turkey Hill Team
164	Congrats to our Eagles and Steelers sweepstakes winners!	2012/2/23	Turkey Hill Team
165	Turkey Hill's newest Limited Edition flavor: Salty Caramel	2012/2/25	Turkey Hill Team
166	Help us determine the final flavors in this year's Ultimate Flavor Tournament!	2012/2/22	Turkey Hill Team
167	Inside the Dairy: Loadin' the trucks and keepin' it cool	2012/2/20	Turkey Hill Team
168	Second Scoop: "Think positive and eat ice cream"	2012/1/25	Turkey Hill Team
169	POLL: Which 2011 Turkey Hill Limited Edition flavors were your favorites?	2011/12/27	Turkey Hill Team
170	Tell us how you cure brain freeze (and our interns will test the cures!)	2012/5/18	Turkey Hill Team
171	Introducing our May Limited Edition Flavor: Southern Lemon Pie	2012/5/23	Turkey Hill Team
172	Happy Strawberries and Cream Day	2012/5/21	Turkey Hill Team
173	Second Scoop: The Turkey Hill Vanilla Ice Cream tasting challenge	2012/5/14	Turkey Hill Team
174	Introducing the Phillies and Moose Tracks Ice Cream Cakes!	2012/5/11	Turkey Hill Team
175	Our Turkey Hill Iced Tea Fan of the Month is…	2012/5/8	Turkey Hill Team
176	Spot the Fake Ice Cream News!	2012/4/23	Turkey Hill Team
177	May poll: Cup or cone?	2012/5/4	Turkey Hill Team
178	Recipe: Ice cream pudding (which flavor will you make?)	2012/4/20	Turkey Hill Team
179	April Limited Edition Flavor: Fried Ice Cream	2012/4/18	Turkey Hill Team
180	View from 30,000 Feet: Delta on Diversity	2012/4/16	Turkey Hill Team
181	recalling nixon in china,40 years later	2012/2/21	JOHN F. BURNS
182	Palestinian's Hunger Strike Challenges Detention by Israel	2012/2/16	ROBERT MACKEY
183	Sleep: why they used to do it twice a night	2012/2/24	Laura Barton

184	Eric Joyce's antics may call time on the cheap parliamentary pint	2012/2/24	Marina Hyde
185	Simple Truths That Help You Save — and Make — Money	2009/9/4	By WSJ Staff
186	Debit Cards Overtake Credit Cards	2009/8/6	Mary Pilon
187	Just as We Learn to Love Saving, Government May Limit 401(k) Contributions	2009/8/31	By WSJ Staff
188	FreeScore.com: More Credit Score Confusion	2009/8/14	Mary Pilon
189	Ahead of New Credit-Card Rules, Expect Higher Rates and More Fees	2009/8/12	Mary Pilon
190	Sorry, there's no such thing as 'correct grammar'	2012/3/2	Michael Rosen
191	Prince Charles goes paintballing – caption competition	2012/3/2	null
192	Sugar addiction is making our children – and Big Food – fat	2012/3/2	Tanya Gold
193	As the recession bites, is a new kind of northern politics emerging?	2012/3/2	Ian Jack
194	Argentinian and British veterans on their Falklands war experiences	2012/3/2	Susanna Rustin
195	From Google downwards, our digital masters must be watched	2012/3/2	Jonathan Freedland
196	American Reporter Marie Colvin’s Final Dispatches From Homs	2012/2/22	ROBERT MACKEY
197	Dutch Puzzled by Santorum’s False Claim of Forced Euthanasia	2012/2/22	ROBERT MACKEY
198	Wounded Journalists Plead for Evacuation From Homs	2012/2/23	ROBERT MACKEY
199	On National Holiday, Australian Leader Forced to Flee Protesters	2012/1/26	J. DAVID GOODMAN
200	Chinese Readers on the ‘iEconomy’	2012/1/25	THE NEW YORK TIMES
201	Celebrate Earth Day With The Nature Conservancy & Delta	2012/4/1	Laura M
202	Delta Flight Simulators Land In Atlanta	2012/4/5	Jeff P
203	Delta Day of Hope: Raising Money For The Fight Against Cancer	2012/5/9	Dory P
204	Six Tips for Your Kid’s Trip	2012/5/31	Jerry F
205	Delta Honor Guard	2012/7/23	Laura M
206	Seder Solutions for Passover	2012/3/25	Elizabeth Smith
207	Berry Time Begins	2012/3/30	James Parker
208	Homemade Remedies For Pet Care	2012/4/29	Elizabeth Smith
209	Putumayo's Latest: African Blues	2012/4/28	Dan Storper
210	A Fresh Take on Fresh Berries Recipe Contest	2012/5/30	Lindsay Robison

211	Fast Forward - Provisioned IOPS for EBS Volumes	2012/8/1	Jeff
212	New High I/O EC2 Instance Type - hi1.4xlarge - 2 TB of SSD-Backed Storage	2012/7/18	Jeff
213	Multiple IP Addresses for EC2 Instances (in a Virtual Private Cloud)	2012/7/6	Jeff
214	CloudFront / Route 53 Edge Location in Sydney, Australia	2012/6/19	Jeff
215	AWS Support - We've Got Your Back	2012/6/14	Jeff
216	Amazon S3 - The First Trillion Objects	2012/6/12	Jeff
217	IAM roles for EC2 instances – Simplified Secure Access to AWS service APIs from EC2	2012/6/11	Anders
218	VM Export Service For Amazon EC2	2012/5/25	Jeff
219	AWS Elastic Beanstalk Now Available in Europe	2012/5/16	Jeff
220	Amazon CloudFront - Support for Dynamic Content	2012/5/14	Jeff
221	Monitor Estimated Charges Using Billing Alerts	2012/5/10	Jeff
222	Amazon RDS for SQL Server and .NET support for AWS Elastic Beanstalk	2012/5/8	Jeff
223	Scalable Session Handling in PHP Using Amazon DynamoDB	2102/4/27	Jeremy
224	Amazon CloudSearch - Start Searching in One Hour for Less Than $100 / Month	2012/4/12	Jeff
225	CloudSpokes Coding Challenge Winners - Build a DynamoDB Demo	2012/4/10	Jeff
226	Amazon S3 - 905 Billion Objects and 650,000 Requests/Second	2012/4/5	Jeff
227	Live Smooth Streaming With CloudFront and Windows IIS Media Services	2012/4/1	Nihar Bihani
228	New Amazon EC2 Fresh Servers	2012/4/1	Jeff
229	Live Streaming with Amazon CloudFront & Flash Media Server 4.5	2012/3/29	Nihar Bihani
230	Updated Amazon Linux AMI (2012.03) Now Available	2012/3/28	Jeff
231	Multi-Region Latency Based Routing now Available for AWS	2012/3/21	Jeff
232	AWS Elastic Beanstalk - Build PHP Apps Using Git-Based Deployment	2012/3/20	Jeff
233	The Next Type of EC2 Status Check: EBS Volume Status	2012/3/12	Jeff
234	EC2 Updates: New Medium Instance, 64-bit Ubiquity, SSH Client	2012/3/7	Jeff
235	Dropping Prices Again-- EC2, RDS, EMR and	2012/3/5	Jeff

	ElastiCache		
236	Calif. E911 Call Confusion, a Deeper Look at the Problem	2012/8/11	Mark J. Fletcher
237	5 Ideas to Balance Business and Customer Experience Objectives	2012/7/10	Laura Bassett
238	Intelligent Call Routing for NG911	2012/7/7	Mark J. Fletcher
239	E911 Apps - The Good - The Bad - and the	2012/6/30	Mark J. Fletcher
240	Shake What Your Mama Gave You!	2012/6/25	Kay Phelps
241	Making E911 Today a Little SMARTER	2012/6/23	Mark J. Fletcher
242	The Great E911 Debate @IAUG Boston	2012/6/4	Mark J. Fletcher
243	Critical Success Factors for Home Agents	2012/6/4	Kay Phelps
244	US School E911 VoIP Call Routed to Canada	2012/5/5	Mark J. Fletcher
245	The Systemic Technological Stagnation of the Emergency Network	2012/4/29	Mark J. Fletcher
246	Celebrating Earth Day	2012/4/20	Wilson Korol
247	New Michigan Law Takes Aim on SWATTERS	2012/3/17	Mark J. Fletcher
248	12 Communication Trends -Worth your attention and planned action now	2012/2/22	Diane Royer
249	Happy 44th Birthday 911!	2012/2/16	Mark J. Fletcher
250	Agile development with Flare	2012/2/9	Bryan Katz
251	Introducing a New Avaya Services blog - Royer's Services Edge: It's all about the client experience	2012/2/7	Diane Royer
252	Avaya Solutions Calculator Adds Networking Analysis	2012/2/6	Wilson Korol
253	E911 - There's an App for that - MAYBE	2012/2/3	Mark J. Fletcher
254	#CyberCivics	2012/1/4	Guy Clinch
255	Trendy Trends for 2012 1 Trendy Trends for 2012	2011/12/13	Rob Daleman
256	Can your Business use a "Zero Email" Policy?	2011/12/7	Jon Alperin
257	Today is a good day: Power of We	2011/12/2	Wilson Korol
258	I Write to Give Thanks	2011/11/27	Guy Clinch
259	The Great E911 Debate 2011, the winner is . . .	2011/11/14	Mark J. Fletcher
260	National EAS test failure a great success?	2011/11/10	Mark J. Fletcher
261	Persuading Gen Y to buy	2012/8/1	Jodi Housman
262	Establish and maintain your professional and personal brand	2012/6/27	Emilie Valle
263	Olympic Games marketing through the years	2012/6/5	Jennifer Newberg
264	Making national go local	2012/4/13	Lindsay Harrington
265	Check your ego at the door	2012/3/28	Regan Dillon
266	Can media relations save lives?	2012/3/9	Maureen O'Connell
267	Timeline offers design and content opportunities	2012/3/1	Mike Hollywood，

	for brands		Alex Nicholson
268	Alex Nicholson	2012/1/20	Jenna Walsh
269	2012 – Observations, Challenges and Predictions	2011/11/1	Cici Gordon
270	The Number One Rule in Event Planning	2011/10/25	Amy Russ
271	Defining Viral: Insights into Facebook's new analytics developments	2011/10/14	Marcus Andrews, Hiroki Murakami
272	The value of focus	2011/9/30	Jessica Anselmi
273	Since when did brands get a personality?	2011/9/23	Mark Malinowski
274	Contests: A do or don't in your marketing plans?	2011/8/25	Amelia Ott
275	Foursquare Event Check-ins: Adding a "What" to the "Where"	2011/8/19	Kathryn Robinson
276	Top 5 takeaways from BlogHer 2011	2011/8/12	Emilie Valle,Jessica Benjamin
277	Keeping luxury luxurious	2011/7/29	Regan Dillon
278	Coupons drive mom word of mouth	2011/7/22	Stephanie
279	The changing news landscape	2011/6/9	Maureen O'Connell
280	Facebook vs. Google: No one wins	2011/5/13	Mike Hollywood
281	I want you to "like" me	2011/4/20	Lindsay Harrington
282	Young consumers key to survival for old brands	2011/4/15	Kimberly Litchfield
283	Branded entertainment: Marketers' new best friend	2011/4/4	Jodi Housman
284	Move fast and break things	2011/3/22	Mike Hollywood, Alex Nicholson, Marcus Andrews
285	What's in a name: Taco Bell manages beef crisis	2011/2/11	Emilee Ellison
286	World IPv6 Day Gets a Dose of Realism with 500 Gbps Firewall Test	2012/6/6	Kyle Flaherty
287	Deep Packet Inspection, Viber, and You	2012/5/3	Andrew Jean
288	Ready for DNSSEC?	2012/3/28	Eric Reeves
289	IPv6 and Carrier-Grade NAT: No Easy Transition	2012/8/6	Eric Reeves
290	Getting Real About VoLTE Testing	2012/7/5	Mike Hamilton
291	Not Even Cryptographically Signed Data Can Be Trusted	2012/6/5	Frank Gifford
292	The Evolution of Testing Standards: Part II	2012/5/29	Mike Hamilton
293	The Power of TCL: Implementing Cisco CLI by Object-Stack	2012/5/23	Jin Qian
294	Protected Against Duqu? Find Out for Sure	2012/3/27	Jarrod Johnson
295	Warning: Slow (TCP) Traffic Ahead	2012/3/1	Scott Register
296	Staying Current, Staying Intelligent: ATI by the Numbers	2012/2/28	Steve McGregory
297	Testing Standards Part I: The Creation and	2012/2/21	Mike Hamilton

	Subsequent Misuse of RFC 2544		
298	Stopping Mobile Malware in the Stack	2012/2/16	Andrew Jean
299	3 Things You Need to Know About Carrier-Grade NAT	2012/2/14	Chuck McAuley
300	Crossbeam Mobile Security Test Was Neither Mobile Nor Security. Discuss.	2012/1/14	Scott Register
301	2011 Blog Rewind: Cyber Range Deployment	2011/12/29	Kyle Flaherty
302	SCADA Attack on City Water Station: What Really Happened?	2011/12/6	Martha Aviles
303	Happy Holidays: 5 Ways to Use DoS Testing to Thwart Cyber Extortion Happy Holidays: 5 Ways to Use DoS Testing to Thwart Cyber Extortion	2011/11/17	Pam O'Neal
304	Live Webcast: Next-Gen Firewall/IPS Testing	2011/11/4	Evan Guthrie
305	Mobile Malware Threats to Enterprise Network Security: Analyst Q&A with VisionGain	2011/10/27	Kyle Flaherty
306	Converged Network Traffic: Why You Should Test Wireline Networks with Mobile Traffic	2011/10/19	Chris Adams
307	The CISO's Guide to Reducing Mobile Malware Threats, Part II	2011/10/15	Kyle Flaherty
308	New Mobile Malware and SCADA Testing Updates to BreakingPoint Products	2011/9/25	Pam O'Neal
309	Vulnerable to Data Breaches? Five Tips for Data Breach Security Testing	2011/8/31	Martha Aviles
310	Router Testing with BGP and Application Traffic: Part 1	2011/8/25	Eric Reeves
311	The Record Price Of The Home Will Be Proportionate Many Factors	2012/6/27	admin
312	Kids to Think and Grow Rich	2012/5/20	admin
313	Whether Or Not An Personal Is Promoting Real Estate	2012/5/17	admin
314	Opt To Get A Stun Gun.	2012/5/8	admin
315	One Of The Leading Manufacturers In The Production	2012/4/18	admin
316	A Low Testosterone, Several Undesirable Effects In The Body	2012/4/15	admin
317	Looking For Big Discount Coach Bags	2012/4/13	admin
318	Many Amazing Trends This Season	2012/4/12	admin
319	Hair Accident	2012/3/20	admin
320	Option Of Scurity Hidden Clock Cameras	2012/2/25	admin
321	It Is A Profession Of Property Management	2012/2/20	admin
322	E-Books	2012/2/14	admin
323	Reverse Assimilation Of The Assembly Process	2012/2/11	admin

324	What Is the Best Online Aggregation	2012/2/8	admin
325	Starting an Acceptation – Consign Search	2012/2/8	admin
326	Wheat And Weeds – Maintain The Garden	2012/2/6	admin
327	Best Admonition For Success Copywriting	2012/2/6	admin
328	Professional Plumbers	2012/2/4	admin
329	Mexican Adobe Pottery	2012/2/4	admin
330	Canopies – Styles And Applications For Access And Walkways	2012/2/3	admin
331	International Affective Companies Or Mover's All-embracing Shipping	2012/1/26	admin
332	By Accepting To Apperceive Your Accessories Options	2012/1/10	admin
333	Dog Training Secrets	2011/12/15	admin
334	Saving live and work	2011/12/3	admin
335	Preparing to let a property	2011/11/29	admin
336	Removing grayish soap residue from white washcloths	2012/8/2	Dr. Laundry
337	Sour smelling colored towels	2012/7/13	Dr. Laundry
338	Sanitize plastic pool toys	2012/7/4	Dr. Laundry
339	Disinfecting a wading pool	2012/7/2	Dr. Laundry
340	How to properly disinfect laundry	2012/6/25	Dr. Laundry
341	How to remove mustard stains	2012/6/1	Dr. Laundry
342	How to restore clothes after a dye transfer	2012/5/29	Dr. Laundry
343	How to remove color bleeding during laundering	2012/5/21	Dr. Laundry
344	Strawberry Stain Removal	2012/5/16	Dr. Laundry
345	Color Bleeding Removal	2012/4/17	Dr. Laundry
346	Old Pet Stain	2012/4/13	Dr. Laundry
347	Smelly Beach Towel	2012/4/4	Dr. Laundry
348	One-Stop-Shop Laundry Products	2012/3/23	Dr. Laundry
349	Gravy Stains	2012/3/21	Dr. Laundry
350	What Does "Non-Chlorine Bleach Only" Mean?	2012/2/10	Dr. Laundry
351	Sharpie® Stain	2012/1/18	Dr. Laundry
352	Underarm Stains	2012/1/16	Dr. Laundry
353	Stained Pillow Cases from Oily Hair	2011/10/7	Dr. Laundry
354	Cleaning White Spandex	2011/8/19	Dr. Laundry
355	Bleach and Colored Clothing	2011/7/21	Dr. Laundry
356	Removing Blood from Clothing	2011/7/20	Dr. Laundry
357	Eradicating Underarm Stains	2011/6/17	Dr. Laundry
358	MRSA Prevention	2011/5/26	Dr. Laundry
359	Ammonia as a Sanitizing Alternative	2011/5/24	Dr. Laundry
360	National Pet Month	2011/5/11	Dr. Laundry
361	Why Groupon and LivingSocial are Flexing their M-Commerce Muscle	2012/8/23	Andrew Lipsman

362	Romney's Online Advertising Ramps Up as Obama Ads Tick Down in July	2012/8/20	Andrew Lipsman
363	Digital Scarcity Improves the Economics of Online Advertising	2012/8/16	Magid Abraham
364	Small Screens for Big Games: How Mobile Phones and Tablets are Influencing the World's Biggest Sporting Event	2012/8/10	Sarah Radwanick
365	Getting the Most out of the Games: How Campaign Optimization Can Drive ROI	2012/8/7	Andrea Vollman
366	MRC Accredits comScore vCE Validation, Including Cross-Domain iFrame Measurement	2012/8/2	Josh Chasin
367	How Search is Helping Quora Break Through to the Mass Market	2012/7/26	Eli Goodman
368	The Digital Battleground: Obama vs. Romney	2012/7/16	Eli Goodman
369	When the Cookie Crumbles	2012/7/2	Gian Fulgoni
370	5 Years Later: A Look Back at the Rise of the iPhone	2012/6/29	Sarah Radwanick
371	The Power of Like 2 Offers New Insights into How Social Marketing Delivers ROI	2012/6/12	Andrew Lipsman
372	Instagram for Video: Searching for the Next Big Thing	2012/6/8	Eli Goodman
373	It's Time to Change the Discussion on Measuring Facebook Effectiveness	2012/6/7	Andrew Lipsman
374	Facebook: Around the World in 800 Days	2012/5/17	Berit Block
375	Social Media Delivers Valuable Exposure for Presidential Campaigns	2012/5/10	Carmela Aquino
376	Despite its Declining Importance for Many Display Campaigns, the Click Remains Important for Pharma Marketers	2012/5/8	John Mangano
377	comScore Introduces Mobile Metrix 2.0: The Next Generation of Mobile Behavioral Measurement	2012/5/7	Mark Donovan
378	Digital Advertising Develops Presidential Candidate Brands and Engages Voter Groups	2012/5/1	Andrew Lipsman
379	Web TV Shows Promise, But Lack of Paid Search Strategy is Puzzling	2012/4/30	Eli Goodman
380	comScore Plants 3 Millionth Tree as Part of "Trees for Knowledge" Program	2012/4/23	Magid Abraham
381	Panel Participation Yields Digital Insights	2012/4/18	Richard Weaver
382	So What if Only 1% of Facebook Fans 'Engage'?	2012/4/4	Andrew Lipsman
383	Optimizing News for Searchers, Search Engines & News Publishers	2012/4/2	Eli Goodman
384	Why businesses are wrong about social media	2012/8/23	Liesel Olson
385	Farm tours offer a closer look into a farmer's	2012/8/20	Elizabeth

	community relationships		Hockerman
386	Why RoboCop and Barbra Streisand are imperative to social media issues management	2012/8/14	Rick Stoner
387	Undergrads, listen up: Here's how to score your dream internship	2012/8/10	Ashley Schaible
388	Take me there: Pure Michigan leaves me wanting more	2012/8/6	Janice Kayser
389	If you want to be creative, stop thinking	2012/7/31	Mike Fredrick
390	Five best practices for a successful Facebook ad campaign	2012/7/27	Adam Engels
391	Mix online and offline for engagement	2012/7/24	Alexa Ganos
392	Tips for pitching broadcast media	2012/7/19	Cassandra Koehler
393	Let me walk a mile in his field	2012/7/17	Pam Golden
394	LinkedIn: Beyond the basics	2012/7/11	Azure Ryerson
395	5 apps that celebrate the Fourth of July	2012/7/3	Olivia Johnson
396	The great Wisconsin duo: Beer and cheese	2012/6/29	Olivia Johnson
397	Innovating beyond technology	2012/6/27	Jon Baade
398	How technology is shaping the dairy industry	2012/6/22	Mallory Steinberg
399	The truths and myths about B2B social media	2012/6/21	Bader Rutter
400	Three signs that you're in a great PR internship	2012/6/20	Bader Rutter
401	What June Dairy Month means to us	2012/6/15	Jen Staudinger
402	How to find success with word-of-mouth marketing	2012/6/14	Grant Thekan
403	"The Pitch" vs. the reality of new business	2012/6/12	Allison Lauer
404	5 ways to celebrate June Dairy Month in "Milkwaukee!"	2012/6/8	Cheryl O'Brien |
405	B2B marketers "grow" at BMA annual conference	2012/6/6	Greg Nickerson
406	Bader Rutter is a Top Workplace	2012/6/5	Jeff Young
407	Motivate your audience by targeting them in new ways	2012/6/1	Damon Sanchez
408	What the Media Wants from Brands	2012/8/29	Susan Gunelius
409	Computer Anxiety & Older Adults: Using Self-Efficacy to Break Usage Barriers	2012/8/29	Erica Harrison
410	What's Next for Branded Apps? [Infographic]	2012/8/29	Susan Gunelius
411	Culture Counts: How to Create a Global User Interface	2012/8/27	Erica Harrison
412	What Consumers Want vs. What Executives Think Consumers Want [Infographic]	2012/8/25	Susan Gunelius
413	The Intersection of Marketing Theory and Investor Relations	2012/8/24	John Palizza
414	PPC Accounts for Only 6% of U.K. Search Clicks	2012/8/24	Susan Gunelius
415	Beautiful or Usable – What Really Matters in	2012/8/22	Erica Harrison

	Corporate Web Design?		
416	4 out of 10 Mobile Device Users Have Clicked on Mobile Ads	2012/8/21	Susan Gunelius
417	Connecting Through the User Interface: 3 Things You Must Know	2012/8/20	Erica Harrison
418	The Untapped Value of the Social Economy	2012/8/17	Susan Gunelius
419	Efficient Markets – Even More on Financial Theory and Investor Relations	2012/8/17	John Palizza
420	You're 50 Milliseconds Away from a Good First Impression	2012/8/16	John Palizza
421	2016 Rio Olympics Logo – An Improvement Over London 2012	2012/8/16	Susan Gunelius
422	Brands Should Be Connecting with Moms on Blogs	2012/8/15	Susan Gunelius
423	Corporate Governance Essentials	2012/8/14	Ed Konczal
424	Using A Theoretical Model to Think About the Value of Investor Relations	2012/8/13	John Palizza
425	New Logo and Colors for Air New Zealand	2012/8/10	Susan Gunelius
426	Extending the J.K. Rowling and Harry Potter Brands	2012/8/8	Susan Gunelius
427	How Well Do Consumers Know the Olympics Brand?	2012/8/6	Susan Gunelius
428	Brands Sponsor Athletes from Great Britain at 2012 Olympics [Infographic]	2012/8/4	Susan Gunelius
429	Nearly 3 out of 4 Online Adults in the U.S. Use Video Sharing Sites	2012/8/2	Susan Gunelius
430	A Penny per Gram	2012/8/13	Dr. Ron Lasky
431	Get SMT Process Engineer Certified @ SMTAI!	2012/7/24	Dr. Ron Lasky
432	Improved Solder Density and Solder Paste Volume Fraction Tool	2012/7/10	Dr. Ron Lasky
433	Cleaning No-Clean Solder Paste Residues in SMT Assembly	2012/7/2	Dr. Ron Lasky
434	Musings on Electronics Assembly Metals: Gold Unified a Country	2012/6/25	Dr. Ron Lasky
435	Good Reads	2012/6/18	Dr. Ron Lasky
436	Lead-Free Electronics Assembly Using the SMT Process - at PCM	2012/5/30	Dr. Ron Lasky
437	Profitability Potential in Printed Circuit Board Assembly	2012/5/15	Dr. Ron Lasky
438	The Miracle of Soldering	2012/4/30	Dr. Ron Lasky
439	Tin Wiskers, Solder Alloy Density, Safety: Sharing a Few Comments	2012/4/16	Dr. Ron Lasky
440	Musings on Metals: Copper	2012/3/28	Dr. Ron Lasky

441	iPad and Kindle Fire? Don't Throw Away Your PC Yet	2012/3/20	Dr. Ron Lasky
442	Cost of Ownership vs Profit Potential In Electronics Assembly	2012/3/7	Dr. Ron Lasky
443	Patty Presents Her Electronics Assembly Copy Exactly Strategy	2012/2/20	Dr. Ron Lasky
444	Is Lowest "Cost of Ownership" the Best Electronics Assembly Manufacturing Metric?	2012/2/14	Dr. Ron Lasky
445	Electronics with 40 Year Shelf Life	2012/2/2	Dr. Ron Lasky
446	Patty Cost Analyzes an MRI scan	2012/1/20	Dr. Ron Lasky
447	Patty Pitches NMAC/I/O	2012/1/13	Dr. Ron Lasky
448	End of the Year Thoughts	2012/1/3	Dr. Ron Lasky
449	Military/Aerospace Lead-Free Solder Reliability Still Unproven	2011/12/12	Dr. Ron Lasky
450	It's a Wrap	2012/8/31	Tami
451	Fused Balsa Wood	2012/8/30	Tami
452	Time to Go Shopping!	2012/8/30	Tania
453	Handsewn birthday napkins	2012/8/29	Emma
454	Twine balls	2012/8/29	Emma
455	Fuse on CraftTV: Episode 2	2012/8/28	Tania
456	Fiskars Fuse Setting in Select Michael's Stores	2012/8/28	News from Fiskars
457	7 more days.....	2012/8/27	Tania
458	Fuse Tech Talk - Coining	2012/8/27	Rebecca
459	YOU did it!	2012/8/27	Tania
460	There's that frame.....AGAIN!	2012/8/26	Tania
461	Last weekend for Home Decor Event Challenges	2012/8/25	Emma
462	A 'Love' Luggage tag	2012/8/23	Emma
463	Are YOU ready to head back to school!?	2012/8/23	Tania
464	Fiskateer Games - Home Decor Event - Challenge #4	2012/8/22	Emma
465	Feeling Challenged?	2012/8/22	Tania
466	Spellbinders Dies in the Fuse Creativity System	2012/8/22	Tami
467	Fuse-ology #5 -Introducing the Adapter Kit	2012/8/21	Tami
468	Fiskateer Games - Home Decor Event - Challenge #1	2012/8/20	Emma
469	Back to School Cuties.	2012/8/20	Tania
470	Fiskateer Games - Home Decor Event	2012/8/20	Emma
471	Who's Ready to ROCK the Home Decor Challenges?	2012/8/19	Tania
472	Warning!! Unauthorized Post!	2012/8/19	Rebecca
473	Glitter!	2012/8/17	Tami
474	Fuse at Michaels	2012/8/17	Tami
475	Magic card TUTORIAL	2012/8/17	Emma

476	INSPIRATION: Cardmaking Challenge #4	2012/8/17	Tania
477	INSPIRATION: Cardmaking Challenge #3	2012/8/16	Tania
478	What I've been doing instead of unpacking...	2012/8/15	Emma
479	'The Hobbit' News: Third Film Sets 2014 Release Date; Second Film Changes Title	2012/8/31	Sean O'Connell
480	Trailer: Michael Shannon is a Killer for Hire in 'The Iceman'	2012/8/31	Derrick Deane
481	Sony Green Lights Will Smith Biblical Vampire Movie	2012/8/31	Derrick Deane
482	Indie Close-Up: Jessica Biel, Jet Li and Some Phone Sex Highlight Indie Films This Weekend	2012/8/31	Derrick Deane
483	Sony Unloads a Double Shot of Featurettes, Check Out 'Dredd 3D' and 'Resident Evil: Retribution'	2012/8/31	Derrick Deane
484	Sound Off: Should Movie Theaters Be Able to Silence Your Phones for You?	2012/8/31	Erik Davis
485	'The Avengers' Returns to Theaters Over Labor Day Weekend	2012/8/31	Derrick Deane
486	EXCLUSIVE: Philip Seymour Hoffman Speaks the Truth in 'The Master' Clip	2012/8/31	Derrick Deane
487	'Taken 2': Liam Neeson Talks About his 'Livid Robot' Character, Plus New Pics and Poster	2012/8/31	Derrick Deane
488	Clint Eastwood's 5 Most Memorable Performances (and No, the RNC Doesn't Count)	2012/8/31	Sean O'Connell
489	Posters: Three New 'Anna Karenina' Character Posters	2012/8/31	Derrick Deane
490	Alamo Drafthouse Guide (8/31-9/5): Hunt a 'Predator' and Attend a High School Reunion	2012/8/31	Brian Salisbury
491	'God of War' and 'Metal Gear Solid' News Show Hollywood's Increased Interest in Video Games	2012/8/31	Sean O'Connell
492	'The Avengers' Deleted Scene Shows Captain America Coping With His Past	2012/8/30	Sean O'Connell
493	Film Mom: Why 'Oogieloves' Shouldn't Scare You, Parents	2012/8/29	Tara McNamara
494	Rachel Weisz, Jessica Chastain, Michael Sheen, Barry Pepper Cut From Terrence Malick's 'To the Wonder'	2012/8/29	Derrick Deane
495	Liam Neeson Will Hunt You Down in New 'Taken 2' Pics	2012/8/29	Derrick Deane
496	Michael Keaton Joins 'RoboCop' Remake	2012/8/29	Derrick Deane
497	You Pick the Box Office Winner: Will ‘The Possession’ Tackle ‘The Expendables 2’?	2012/8/29	Erik Davis
498	Trailer: Jackie Chan Dons Rollerblade Suit in	2012/8/29	Derrick Deane

	'Chinese Zodiac'		
499	Watch: Amy Adams Gets Angry in 'The Master', Anna Kendrick Goes A Capella in 'Pitch Perfect'	2012/8/29	Derrick Deane
500	Ryan Gosling To Make Directorial Debut With Film Starring Christina Hendricks	2012/8/29	Derrick Deane
501	Spike Lee's Michael Jackson Documentary to Air on ABC During Thanksgiving	2012/8/29	Derrick Deane
502	Rumor: 'Before Sunset' Follow-Up Currently Filming in Greece	2012/8/29	Derrick Deane
503	'Bull Durham' Writer Ron Shelton Returns to Sports Films, Has Written Officially Licensed NFL-Themed Movie	2012/8/29	Derrick Deane
504	Exclusive: 'The Oranges' Poster Premiere!	2012/8/29	John Halecky
505	Batman Should Retire Until the 'Justice League' Movie Opens in Theaters: Discuss	2012/8/29	Sean O'Connell
506	'The Avengers,' 'Beverly Hills Cop' are Latest Movies Moving to the TV Screen	2012/8/29	Sean O'Connell
507	Daily Recap: Sam Raimi Blocks 'Evil Dead 4' From Happening, Talks Up 'Wildly Bloody' Remake	2012/8/28	Derrick Deane
508	Bill Condon Talks 'Breaking Dawn' Runtime, Renesmee and 'The Biggest Musical Number I've Ever Directed'	2012/8/28	Derrick Deane
509	Beat this caption: Asleep at the cone	2012/8/30	Turkey Hill Team
510	Quotes rewritten: How would you spend some extra money?	2012/8/29	Turkey Hill Team
511	Do you remember the first time you tried Turkey Hill Ice Cream?	2012/8/28	Turkey Hill Team
512	I scream, you scream, we all scream for funny ice cream videos	2012/8/24	Turkey Hill Team
513	Chocolate addiction? We've got the remedy!	2012/8/23	Turkey Hill Team
514	Inside the Dairy: Enter Here!	2012/8/20	Turkey Hill Team
515	Second Scoop: Ahh, to be young again!	2012/8/17	Turkey Hill Team
516	Our August Iced Tea Fan of the Month is…	2012/8/14	Turkey Hill Team
517	Ask Ernie: What would you be doing if you weren't making ice cream?	2012/8/13	Turkey Hill Team
518	Anyone in the mood for a peanut butter banana milkshake?	2012/8/10	Turkey Hill Team
519	Solar-powered ice cream cart turns sun into delicious fun!	2012/8/8	Turkey Hill Team
520	A sneak peek at our 2013 Limited Edition ice cream line-up	2012/8/6	Turkey Hill Team
521	August prize: The ice cream necklace!	2012/6/3	Turkey Hill Team

522	Are you ready for the next Cow Patrol?	2012/8/2	Turkey Hill Team
523	The Turkey Hill interns test eight brain freeze cures!	2012/8/1	Turkey Hill Team
524	After Pussy Riot Verdict, Christian Culture Warriors Run Riot in Moscow	2012/8/30	ROBERT MACKEY
525	Latest Storm Updates on Isaac	2012/8/29	CHRISTINE HAUSER
526	Witness to Rachel Corrie's Death Responds to Israeli Court Ruling Absolving Soldier	2012/8/28	ROBERT MACKEY
527	Updates on Hurricane Isaac	2012/8/28	MICHAEL SCHWIRTZ
528	Morsi's Syria Plan Suggests Regional Approach to Foreign Affairs	2012/8/27	DAVID D. KIRKPATRICK
529	Multiple Clips of Syrian Helicopter Crash	2012/8/27	ROBERT MACKEY
530	After Deadly Sweep Through Haiti, Tropical Storm Isaac Heads for Gulf Coast	2012/8/27	CHRISTINE HAUSER and JENNIFER PRESTON
531	Colleagues and Stargazers Hail Armstrong After Death	2012/8/25	PATRICK MCGEEHAN
532	Moscow Court Finds Kasparov Not Guilty of Illegal Protest During Pussy Riot Trial	2012/8/24	ILYA MOUZYKANTSKII and ROBERT MACKEY
533	Work at Slaughterhouse Is Halted After Graphic Undercover Videos	2012/8/22	CHRISTINE HAUSER
534	Japanese Journalist's Final Report From Syria Is Released After Her Death	2012/8/22	ROBERT MACKEY
535	Anti-Islam Ads Remixed in San Francisco and New York	2012/8/21	ROBERT MACKEY
536	United States Grants Broad Sanctions Exemption for Iran Quake Aid	2012/8/21	RICK GLADSTONE
537	How to Customize the Map View of Garmin's StreetPilot Onboard App for iPhone	2012/8/16	null
538	Peg's Posts: Girls on the Run founder Molly Barker comes to Garmin	2012/8/16	null
539	Lucy's Logbook: Transferring routes to the Garmin Pilot app	2012/8/15	null
540	White Marlin Open - winner takes home another Garmin Grand	2012/8/13	null
541	Team Garmin: Team win in Utah puts Vande Velde in leader's jersey	2012/8/9	null
542	Lucy's Logbook: Tips for connecting the GDL 39	2012/8/8	null

	to your device		
543	Garmin selected by Maritimo at the Sydney Boat Show	2012/8/8	null
544	Peg’s Posts: Trade in your old Forerunner, get $50 back on a new one	2012/8/7	null
545	NAVIGON Android App Update Includes 3D Map Views and Last Mile Navigation	2012/8/2	null
546	Motivate the next generation: join SoleMates for Girls on the Run	2012/8/2	null
547	Lucy's Logbook: Ensure database accuracy with Aviation Data Alerts	2012/8/1	null
548	Garmin GPS 19x 10Hz antenna in action	2012/7/31	null
549	Garmin and Tri-Tronics Throwdown: Track and Train Your Hunting Dogs with Alpha	2012/7/31	null
550	Garmin Triathlete Sunita takes on Ironman Racine 70.3	2012/7/30	null
551	Traveling to London? Save now on the Garmin and NAVIGON iPhone navigation apps	2012/7/30	null
552	Oshkosh Day 5: AirVenture Photo Album	2012/7/27	null
553	Fishing the Gulf with "Deadliest Catch" Deckboss, Edgar Hansen	2012/7/27	null
554	Jake's Journal: This summer, explore London like a local	2012/7/26	null
555	Oshkosh Day 4: GDL 39 FAQs	2012/7/26	null
556	Team Garmin: Candid Q&A with Christian Vande Velde	2012/7/25	null
557	Oshkosh Day 3: Vans RV 40th anniversary, New affordable database option	2012/7/25	null
558	Storage Switzerland Supports Vertical Integration for SSD/HDD Price Parity	2012/9/14	Hu Yoshida
559	What defines an enterprise storage system?	2012/9/11	Hu Yoshida
560	Conversation with Dave Wilson on the State of IT in the Health Care Industry	2012/9/7	Hu Yoshida
561	Why build your own SSD Controller?	2012/9/5	Hu Yoshida
562	Penske-Hitachi Car Wins August 26 Indy Race in Sonoma, California	2012/8/31	Hu Yoshida
563	Is 1 Million IOPS the new Threshold for Enterprise Storage?	2012/8/30	Hu Yoshida
564	Making Flash Ready for Enterprise Primetime	2012/8/23	Hu Yoshida
565	Media and Entertainment: Unleash Creativity with Hitachi NAS Platform	2012/8/20	Hu Yoshida
566	Storage Technology Certification	2012/8/14	Hu Yoshida
567	Redefining Unified Storage – Unified	2012/8/10	Hu Yoshida

	Management		
568	Redefining Unified Storage – HUS Object based File System	2012/8/8	Hu Yoshida
569	Redefining Unified Storage – Hardware Differences Part 2	2012/7/25	Hu Yoshida
570	Hybrid Storage Controllers	2012/7/19	Hu Yoshida
571	Redefining Unified Storage – Hardware Differences Part 1	2012/7/16	Hu Yoshida
572	Redefining Unified Storage	2012/7/6	Hu Yoshida
573	Hitachi Content Platform – Object Storage for Standard Applications	2012/6/28	Hu Yoshida
574	Social Innovation	2012/6/27	Hu Yoshida
575	Top Workplaces in Silicon Valley and Bay Area	2012/6/21	Hu Yoshida
576	Video Surveillance for the Next Generation	2012/6/18	Hu Yoshida
577	What is a good target for capacity utilization?	2012/6/13	Hu Yoshida
578	Here we go – another fight	2012/9/24	IAG
579	UK survey shows the way for IFE	2012/9/21	IAG
580	Lufthansa dealing with LCCs	2012/9/20	IAG
581	FastJet confirms airbase in Tanzania	2012/9/19	IAG
582	Aer Lingus – bolshie labor again	2012/9/14	IAG
583	Emirates continues its US expansion	2012/9/13	IAG
584	Apple's unexpected advance into aviation	2012/9/12	IAG
585	The coming LCC war in Asia	2012/9/11	IAG
586	Middle East growth kicks in	2012/9/10	IAG
587	The iPad continues to conquer	2012/9/7	IAG
588	Qantas and Emirates – the new aviation world	2012/9/6	IAG
589	Stupid is what stupid does (part of the series)	2012/9/5	IAG
590	India continues to supply too much aviation humor	2012/9/4	IAG
591	Canada approves GoGo	2012/8/31	IAG
592	Airbus develops China biofuels opportunity	2012/8/30	IAG
593	Here we go – EU investigates Ryanair's interest in Aer Lingus	2012/8/29	IAG
594	Airbus' huge win at PAL	2012/8/28	IAG
595	Qantas and the 787-9	2012/8/24	IAG
596	Virgin to start domestic UK service	2012/8/21	IAG
597	AMR labor request denied	2012/8/16	IAG
598	Holiday season coming early?	2012/8/14	IAG
599	Another bites the dust	2012/8/13	IAG
600	100-149 Seat Market is Viable	2012/8/10	IAG
601	VJ's letter	2012/8/9	IAG
602	Sukhoi in trouble?	2012/8/8	IAG
603	Boeing's continued acceleration	2012/8/6	IAG
604	Airline meals give the needle	2012/8/3	IAG

605	London's "new" airport	2012/8/2	IAG
606	Qantas & Emirates	2012/8/1	IAG
607	Comair and the 50-seat RJ	2012/7/31	IAG
608	Skip the Lines! Here's How to Get a Free iPhone 5 …	2012/9/20	Joseph Manna
609	September's 'Northern Lights' Desktop Wallpaper	2012/9/6	Joseph Manna
610	Webinar Series: The Path to the Inbox	2012/8/30	Joseph Manna
611	Let's Help Iron Tribe Fitness Speak at SXSW 2013!	2012/8/22	Joseph Manna
612	Hidden Product Insights from Your Customers	2012/8/13	Joseph Manna
613	INFOGRAPHIC: Four Ways to Boost Business with CustomerHub	2012/8/7	Joseph Manna
614	August's 'Boat Project' Desktop Wallpaper	2012/8/6	Joseph Manna
615	Customer Hub Integration with Infusionsoft Provides Maximum Automation	2012/7/26	Kyle Leavitt
616	Vote for Your Favorite Small Business Influencer	2012/7/25	Joseph Manna
617	Why I Love Small Businesses	2012/7/15	Brian Gates
618	User Groups Are Happening in a City Near You	2012/7/15	Joseph Manna
619	Developers: Become the Champion in our 'Battle of the Apps' Contest	2012/7/12	Ammon Curtis
620	Small Business Success Club Helps Entrepreneurs Grow	2012/7/11	Joseph Manna
621	4 Insanely Simple Ways to Protect Your Twitter Account	2012/7/6	Joseph Manna
622	July's 'Summer Picnic' Desktop Wallpaper	2012/7/5	Joseph Manna
623	Introducing the Big Ideas Blog	2012/6/28	Joseph Manna
624	Consistent Pursuit of a Worthy Goal	2012/6/25	Madison Jacobs
625	An Insider's Guide to Write Powerful & Simple Emails	2012/6/19	Guest Blogger
626	Turn Your Customers into Friends: Why Facebook for Small Business is Personal	2012/6/18	Guest Blogger
627	2012 Small Business Success Tour Kicks Off in NYC	2012/6/15	Katie Healy
628	On World Alzheimer's Day, Let's Work to Eliminate Stigma about Dementia	2012/9/21	JNJBTW
629	Finding Support When Facing Serious Illness	2012/9/12	JNJBTW
630	Janssen Featured at White House World Hepatitis Day Activities	2012/9/7	JNJBTW
631	Greening our Hospitals Will Reduce Healthcare Costs Now and in the Future	2012/9/4	JNJBTW
632	IBD Icons – Casey Abrams Inspires Others to Live Beyond Ulcerative Colitis and Crohn's Disease	2012/8/27	JNJBTW
633	August, the Month to Raise Awareness about	2012/8/20	JNJBTW

	Psoriasis!		
634	Our Safety and Care Commitment	2012/8/15	JNJBTW
635	Seeing the Differences We Make In Our Communities	2012/8/14	JNJBTW
636	BlogHer 2012: Social media for social good	2012/8/9	JNJBTW
637	Johnson & Johnson at BlogHer 2012 – Come and See Us There!	2012/8/2	JNJBTW
638	Fighting HIV/AIDS, Part of Our Commitment to Caring	2012/7/22	JNJBTW
639	New ennTV Episode: Meet DePuy Synthes!	2012/7/20	JNJBTW
640	2012 Dr. Paul Janssen Award for Biomedical Research Honors Joint Discovery of micro-RNA	2012/6/19	JNJBTW
641	BAND-AID® Brand Magic Vision	2012/6/16	JNJBTW
642	2012 Pacific Health Summit: Collaborating to Deliver Game Changers in Healthcare	2012/6/14	JNJBTW
643	Making Awareness a Priority for African-Americans — Putting Prostate Cancer on the M.A.P.	2012/6/5	JNJBTW
644	New ennTV Episode: Sun Protection, Inspiring Innovation and Fitness Tips!	2012/5/31	JNJBTW
645	A Turning Point in Alzheimer's Disease – Reasons to be Optimistic	2012/5/22	JNJBTW
646	Today is World Day for Diversity	2012/5/21	JNJBTW
647	Greening the Healthcare Supply Chain with Sustainable Purchasing	2012/5/17	JNJBTW
648	CHINA'S "NOBILITY OF NATURE" - BEES, HONEY & CLEAN WATER	2012/5/15	Bill Marriott
649	EXECUSTAY ON THE MOVE	2012/5/21	Bill Marriott
650	AROUND THE WORLD IN A DAY	2012/5/29	Bill Marriott
651	LIVING WELL	2012/6/4	Bill Marriott
652	FIVE YEARS BLOGGING - PEN PALS, PILATES AND HITTING PUBLISH	2012/6/11	Bill Marriott
653	22 HOTELS IN 10 DAYS - MY TRIP TO CHINA	2012/6/22	Bill Marriott
654	BEAUTIFUL BATHROOMS	2012/7/2	Bill Marriott
655	300 MILLION VIEWERS WATCHING ON CCTV	2012/7/9	Bill Marriott
656	MARRIOTT IN THE LIMELIGHT	2012/7/15	Bill Marriott
657	MY COOK-OFF CHALLENGE	2012/7/23	Bill Marriott
658	Isles of Idyll — An Excerpt from "Crossings"	2012/5/14	Michael Kew
659	American Rivers Announces America's Most Endangered River for 2012	2012/5/16	Emily Nuchols
660	Park Rumors	2012/5/17	Kelly Cordes
661	Follow-up to Big Data – First Steps	2012/6/12	Hu Yoshida
662	Hitachi's Integrated Vision Around Big Data	2012/6/8	Hu Yoshida

663	Improving Storage Densities When Areal Density Slows Down	2012/6/6	Hu Yoshida
664	Vote for HDS at VMworld	2012/6/4	Hu Yoshida
665	CB500 With Embedded Switch	2012/5/31	Hu Yoshida
666	The Value of Big Data – First Steps	2012/5/29	Hu Yoshida
667	Application Integration Requires More Than JBODs	2012/5/21	Hu Yoshida
668	Paddleboard World Championships-- A Preview With Jenny Kalmbach	2012/7/24	null
669	Oshkosh Day 2: Garmin Pilot app gets more features, more affordable	2012/7/24	null
670	Another Garmin Grand Giveaway winner!	2012/7/23	null
671	Get Connected and Be Creative with Garmin Chirp Wireless Beacon	2012/7/23	null
672	EAA AIrVenture is underway in Oshkosh	2012/7/23	null
673	How Much Data Do Connected Apps Use While on the Road?	2012/7/20	null
674	Jake's Journal: Ride with Team Garmin and join us in the fight against MS	2012/7/19	null
675	Garmin volunteers put Cub Scouts on path toward Outdoor Activity Award	2012/7/19	null
676	Garmin Connect Rolls Out Personal Records in the First of Several Exciting Summer Updates	2012/7/18	null
677	GWX 70: Affordable Doppler-Capable Weather Radar for General Aviation	2012/7/18	null
678	New Features for GTN Touchscreen Avionics	2012/7/17	null
679	Team Garmin: Vande Velde and argyle armada vow to keep fighting	2012/7/17	null
680	Garmin Leads the Way to ADS-B Future	2012/7/16	null
681	Team Garmin: Millar wins thriller in Stage 12 of Tour	2012/7/13	null
682	GPS and Garmin Help Rescuers, Refugees in Colorado Wildfires	2012/7/13	null
683	Get Lost In The Hunt With Garmin	2012/7/13	null
684	Garmin GSD 26 CHIRP sonar takes on the Atlantic	2012/7/12	null
685	Bad Boys Run; Good Girls Eat Donuts?	2012/7/12	null
686	Scenes from ICAST - the world's largest sportfishing tradeshow	2012/7/11	null
687	Garmin GLO Portable GPS and GLONASS Receiver Brings High-Integrity GPS Capability to Mobile Devices	2012/7/11	null
688	Helping Small Businesses Every Step of the Way	2012/6/14	Madison Jacobs

689	Confusion over Consolidation of Storage and Consolidation of Data	2012/5/18	Hu Yoshida
690	HCP supports Big Data at PB Scale	2012/5/17	Hu Yoshida
691	Big Data Variety	2012/5/14	Hu Yoshida
692	99 Notre Dame, San Jose	2012/5/8	Hu Yoshida
693	Big Data Velocity Requirements	2012/5/4	Hu Yoshida
694	Big Data Volume Requirements	2012/5/2	Hu Yoshida
695	Big Data Origins	2012/4/26	Hu Yoshida
696	The BUZZ about HUS	2012/4/24	Hu Yoshida
697	How many ways are we connected? Let me count the ways	2012/4/16	Hu Yoshida
698	Hitachi Announces Hitachi Compute Blade 500 with Intel Xeon E5-2600 processors and embedded 10 Gig Ethernet Switch	2012/4/12	Hu Yoshida
699	Seagate's HAMR Reaches the 1TB Per-Square-Inch Barrier and HGST Announces 4TB Disk	2012/4/11	Hu Yoshida
700	Reducing Costs or Being More Efficient?	2012/4/10	Hu Yoshida
701	Hitachi NAS Differences for Cloud and Big Data	2012/4/6	Hu Yoshida
702	The Beat Goes On	2012/4/3	Hu Yoshida
703	Are Tier 1 Controllers Relevant?	2012/3/30	Hu Yoshida
704	Pas De Deux: Brocade and Hitachi Data Systems for zEnterprise	2012/3/29	Hu Yoshida
705	Convergence of File, Block and Objects	2012/3/22	Hu Yoshida
706	HDS named World's Most Ethical Company	2012/3/20	Hu Yoshida
707	Storage Management Efficiency –A Unified Approach	2012/3/19	Hu Yoshida
708	Virtualization Redefines Tier 1 Storage	2012/3/14	Hu Yoshida
709	Application Efficiencies	2012/3/12	Hu Yoshida
710	Energy Efficiencies	2012/3/5	Hu Yoshida
711	Information: No Longer for the Privileged Few	2012/3/2	Hu Yoshida
712	Storage Data Protection Efficiencies	2012/2/24	Hu Yoshida
713	Do More With What You Have	2012/2/23	Hu Yoshida
714	Additional Storage Performance Efficiencies for Mainframes	2012/2/21	Hu Yoshida
715	Storage Performance Efficiency	2012/2/17	Hu Yoshida
716	What We're Watching	2012/5/14	Gloria Huang
717	Uniting Symbols – Creating Hope	2012/5/14	Kristiana Almeida
718	It's True: Zombies Love the Red Cross	2012/5/15	Kristiana Almeida
719	Civil War on Sunday	2012/5/16	Erin Ferris
720	Infographic: Amazing Feats of All Star-Blood Donors	2012/5/18	Gloria Huang
721	Weekly Worldwide Wrap-Up	2012/5/18	Nate Warren

722	Gladiator Fire	2012/5/21	Gloria Huang
723	Honoring Joplin	2012/5/22	Gail J McGovern
724	Baby(sitter) on Board	2012/5/23	Gloria Huang
725	National Mental Health Month 2012	2012/5/24	Erin Ferris
726	Voting for Toyota’s 100 Cars for Good Opens This Morning at 10 am ET/7 am PT	2012/5/14	null
727	All-New 2012 Yaris: A Great Blend of Performance, Convenience, MPG and Safety	2012/5/14	null
728	Lexus Announces Voluntary Safety Recall of Certain 2013 Model Year GS 350 F Sport RWD Vehicles	2012/5/16	null
729	Toyota to Increase V6 Engine Production in Alabama	2012/5/17	null
730	Toyota Hits Four Million Unit Global Sales Milestone for Hybrids	2012/5/22	null
731	Toyota Earns Three 'Top Scoring Car Awards' in First Annual Total Car Score Rankings	2012/5/23	null
732	America's Best-Selling Car Tuned Up for 2013 NASCAR Season	2012/5/23	null
733	Scion tC Receives Autobytel and AutoPacific 2012 Vehicle Satisfaction Award	2012/5/24	null
734	Toyota Announces Pricing for 2013 Venza Crossover	2012/5/25	null
735	Toyota Leads the Way in Green Buildings	2012/5/30	null
736	http://blog.logos.com/2012/09/dr-charles-stanley-featured-contributor-to-the-faithlife-study-bible/	2012/9/21	Jayson Bradley
737	Free Books for Your App Insights	2012/9/20	Nathan Smoyer
738	Abraham Ships Soon—Pre-Order It Today!	2012/9/19	John D. Barry
739	4 Reasons the Westminster Bible Companion Is for Everyone	2012/9/18	Brandon Rappuhn
740	Logos 4: Return to a Previous Logos Desktop Display	2012/9/17	Morris Proctor
741	Celebrate Dr. Charles Stanley and Save Big!	2012/9/15	Jayson Bradley
742	The Time Has Come to Apply for Your DMin!	2012/9/13	Jayson Bradley
743	Got a Camera? Win Cool Prizes!	2012/9/13	Jayson Bradley
744	Save on the Updated JPS Commentary Collection	2012/9/12	Deni Avant
745	Only 3 Days Left for Back-to-School Savings!	2012/9/11	Jayson Bradley
746	The Faithlife Study Bible Advantage: Part 1	2012/9/11	Jayson Bradley
747	Logos 4: Collins Thesaurus of the Bible	2012/9/10	Morris Proctor
748	3 Reasons Logos Resources Are So Valuable	2012/9/7	Jayson Bradley
749	Advance Notice on Discontinued Platform Support	2012/9/7	Jayson Bradley
750	Receive Chapters of Reformed Dogmatics Before	2012/9/6	Jayson Bradley

	It's Released!		
751	Back to School Sale: Karl Barth's Church Dogmatics	2012/9/5	Cliff Kvidahl
752	Win an iPad and a Base Package from the Glorious Ruin Tour!	2012/9/4	Nathan Smoyer
753	Logos 4: Power Reading	2012/9/3	Morris Proctor
754	Free Book of the Month: Psalms and Hymns of Isaac Watts	2012/9/1	Jayson Bradley
755	Experience the Biblical Story with Rich Media	2012/8/31	Jayson Bradley
756	An Interview with Joel B. Green, Editor of the NICNT	2012/8/31	Cliff Kvidahl
757	Let Lexham Do Your Research Legwork!	2012/8/30	John D. Barry
758	Better Bible Study on iPad & iPhone—Free!	2012/8/29	Kensey Burdick
759	Stanley Hauerwas: Time Magazine's Choice for America's Best Theologian	2012/8/28	Brandon Rappuhn
760	4 More Days to Download August's Free Book!	2012/8/28	Jayson Bradley
761	Save on Walter Brueggemann While You Still Can!	2012/8/27	Brandon Rappuhn
762	Logos 4: Visual Filter for the Greek Words Translated Temple	2012/8/27	Morris Proctor
763	Win a Scholar's Library from Faithlife	2012/8/24	Jayson Bradley
764	Knox DMin Classes Kick Off in Bellingham	2012/8/24	Jayson Bradley
765	Save Now on Our Most Popular Resources!	2012/8/23	Jayson Bradley
766	Start an online reputation quilting bee	2012/8/23	Chris Abraham
767	A strong offense is the best online reputation defense	2012/8/22	Chris Abraham
768	How to marry your online reputation into the Google Search Index	2012/8/21	Chris Abraham
769	Neuromarketing: The Future of Advertising	2012/8/14	speters
770	Integrity is inherent in earned media	2012/8/12	Chris Abraham
771	Outsource your social media to the professionals	2012/8/11	Chris Abraham
772	Flipboard is changing everything for online content creation	2012/8/10	Chris Abraham
773	reddit is the 800-pound gorillaphant in the room	2012/8/8	Chris Abraham
774	Developing a Strong Social Media Presence	2012/8/7	speters
775	Don't mess with reddit (or Texas)	2012/8/7	Chris Abraham
776	3 tools to help you master the art of unfollowing	2012/8/7	Chris Abraham
777	The Changing World of Marketing	2012/7/26	speters
778	The Strange New World Of Social Media Optimization	2012/7/12	speters
779	Collaboration: A Key to Marketing Success	2012/7/10	speters
780	5 SEO Tips to Optimize Google+	2012/7/3	speters
781	5 Ways to Build Your Online Visibility with	2012/6/18	speters

	Pinterest		
782	Stop shouting and start listening and responding online	2012/6/16	Chris Abraham
783	Keep blogging after you think it’s a stupid waste of your time	2012/6/15	Chris Abraham
784	Never let them catch you rolling your eyes	2012/6/14	Chris Abraham
785	Be a consistent social media parent	2012/6/13	Chris Abraham
786	You need to be social media sluttier	2012/6/12	Chris Abraham
787	MySQL Instance from Generic Binary on Linux	2012/9/26	rajeshr
788	Enterprise Trade Compliance: Changing Trade Operations around the World	2012/9/25	John Murphy
789	Oracle releases new Java Embedded products	2012/9/25	Henrik Stahl
790	Multiple OpenSSL vulnerabilities in Sun SPARC Enterprise M-series XCP Firmware	2012/9/25	RitwikGhoshal
791	Oracle OpenWorld / JavaOne Where I'll Be	2012/9/25	Shay Shmeltzer
792	Oracle Enterprise Manager 12c Anniversary at Open World General Session and Twitter Chat using #em12c on October 2nd	2012/9/25	Anand Akela
793	Business Forecast: Cloudy with a Chance to Gain	2012/9/25	Oracle OpenWorld Blog Team
794	OneGoLive Services Delivers its Oracle Accelerate Solution for Wineries for Oracle’s JD Edwards EnterpriseOne Provides comprehensive ERP solution for wineries	2012/9/25	LanaProut
795	EVRY Delivers Nordic Standard Statutory Application (NSSA), Oracle Accelerate Solution based on Oracle Hyperion Financial Management	2012/9/25	LanaProut
796	National Marrow Donor Program Upgrades to Oracle E-Business Suite 12.1 to Boost Operational Efficiency and Visibility	2012/9/25	LanaProut
797	Absence Management White Papers to Assist with your Implementations	2012/9/25	Carolyn Cozart
798	Meet the Java EE 7 Specification Leads BOF at JavaOne	2012/9/25	Linda DeMichiel
799	Oracle OpenWorld and JavaOne are right around the corner!	2012/9/25	Eric Jensen
800	Oracle releases new Java Embedded products	2012/9/25	Henrik Stahl
801	Discovery methods	2012/9/25	Owen Allen
802	Oracle Solaris Cluster at Oracle OpenWorld 2012	2012/9/25	evek
803	Supercharging the Performance of Your Front-Office Applications @ OOW'12	2012/9/25	Sanjeev Sharma
804	Enterprise Trade Compliance: Changing Trade Operations around the World	2012/9/25	John Murphy

805	General Session: Building and Managing a Private Oracle Java and Middleware Cloud	2012/9/25	Ruma Sanyal
806	CSO Summit @ Executive Edge	2012/9/25	Naresh Persaud
807	How to integrate Java ME SDK 3.2 with Eclipse	2012/9/25	SungmoonCho
808	How to integrate Java ME SDK 3.2 with NetBeans	2012/9/25	SungmoonCho
809	Find Thousands of Oracle Jobs on oDesk	2012/9/25	Brandye Barrington
810	Special Activities in the OTN Lounge	2012/9/25	Bob Rhubart
811	Oracle OpenWorld - Events of Interest	2012/9/25	Larry Wake
812	Oracle Unveils Oracle Java Embedded Suite 7.0	2012/9/25	Darryl Mocek
813	Java ME SDK 3.2 is now live	2012/9/25	SungmoonCho
814	Java Embedded Releases	2012/9/25	Tori Wieldt
815	Organizing Before and After Vacations	null	Rachel Oja
816	Toys, Bows and Legos – Oh My! Kid Organizing Tips	null	Rachel Oja
817	Quick, Easy Laundry Room Transformation	null	Rachel Oja
818	What's Your Clean Cycle?	null	Deb Lee
819	5 Steps to An Organized Closet	null	Beth Zeigler
820	5 Things You Can Purge From Your Home… Right Now	null	Beth Zeigler
821	Letting housecleaning be easy	null	Janine Adams
822	4 Steps to an Organized Fridge	null	Rachel Oja
823	How to Use and Reuse…	null	Rachel Oja
824	Keep or Not Keep: 4 Tips to Making Good Decisions	null	Rachel Oja
825	Make it easy to put stuff away	null	Janine Adams
826	6 Things To Include in a 'No Fuss' Entry	null	Beth Zeigler
827	2 Simple Ways to Find Your Everyday Things	null	Deb Lee
828	Keeping Pieces Together Part of Organizing Puzzle	null	Rachel Oja
829	5 Ways to Make Garage Organization F-U-N!	null	Beth Zeigler
830	Use Bento for Gifts	null	Rachel Oja
831	Office Supplies On The Move	null	Rachel Oja
832	5 Reasons We Clutter & Solutions To Stay Ahead of the Clutter Curve	null	Rachel Oja
833	Six steps to an organized bathroom	null	Janine Adams
834	8 Things to do Right Away for an Organized Basement	null	Rachel Oja
835	When commuting is competitive and fun	2012/9/24	Niall Baker
836	Unity in diversity – the view from an Indian IT hub	2012/9/18	Annalise Toberman
837	Let's ask The Audience ……..	2012/9/14	Kate Anderson
838	Olympic fan wars	2012/9/13	Clayton Fussell

839	The consequences of a misplaced. full stop	2012/9/7	Fiona Pannell
840	The Olympics, nationalism and the power of imagination	2012/8/13	Louise Amantani
841	On stage at the 2012 Olympic opening ceremony	2012/8/6	Sandria Terrelonge
842	Are you spoilt for choice?	2012/7/31	Vanessa Fouscas
843	Y should I?	2012/7/25	Charlotte Crichton
844	Comic Sans – the font of all insight?	2012/7/6	Kate Anderson
845	When things go wrong…..do the right thing!	2012/7/5	Niall Baker
846	RS and Leapfrog, defying the laws of mathematics	2012/7/2	RS Consulting
847	Decision making in the blink of an eye	2012/6/21	Ivonne Hoeger
848	The paradoxical theory of change	2012/6/7	Victoria Boelman
849	Why we don't have strawberry flavoured crisps	2012/5/25	Kate Anderson
850	Behavioural economics to promote healthy behaviour	2012/5/25	Annalise Toberman
851	Cupcakes and vegetables – changing customer ideals	2012/5/21	Kate Downer
852	The revolution will not be televised … It will be segmented	2012/5/18	Kate Anderson
853	Are you talking to me? Or, what Justin Bieber has taught me about trade shows	2012/5/16	Martin Wootton
854	A history of the world – in 140 characters	2012/5/15	Clayton Fussell
855	Attack of the killer journalists from hell	2012/5/9	Andrew Wood
856	The Girls Go Coastal	2012/9/26	Chloe Ochse
857	West Nile Virus and Your Pets	2012/9/21	Chloe Ochse
858	Getting Ready to Hit the Road	2012/9/18	Chloe Ochse
859	Thinking About Adding a New Cat to the Jungle?	2012/9/14	Chloe Ochse
860	A First Time Foster	2012/9/11	Chloe Ochse
861	Five Dogs and a Cat or Two	2012/9/7	Chloe Ochse
862	Comparing Pet Health Insurance Plans	2012/8/31	Chloe Ochse
863	Sleeping Beauty	2012/8/28	Chloe Ochse
864	The Fat You Don't Need	2012/8/24	Chloe Ochse
865	Getting Back to Basics - Doggy Style	2012/8/21	Chloe Ochse
866	The Fat You Need	2012/8/17	Chloe Ochse
867	Tiny Traveling Terrors	2012/8/14	Chloe Ochse
868	The Best Time to Insure a Pet	2012/8/10	Chloe Ochse
869	On the Road Again	2012/8/7	Chloe Ochse
870	Heart Worm Prevention	2012/8/3	Chloe Ochse
871	Indoor Cats Need Wellness Plans Too	2012/7/27	Chloe Ochse
872	Maizie, the Psychic	2012/7/24	Chloe Ochse
873	Canine Influenza; Should You Vaccinate your Dog?	2012/7/20	Chloe Ochse

874	Nothing like Coming Home	2012/7/17	Chloe Ochse
875	The Lowly Dog House Revisited	2012/7/13	Chloe Ochse
876	Discounting to drive up impulse buying	2012/5/3	Kate Anderson
877	6 tips for marketing the cloud	2012/4/20	Martin Wootton
878	18/04/2012 – Processes and costs of transferring a pension scheme	2012/4/17	RS Consulting
879	5 ways to funk up your co-creation	2012/4/17	Victoria Boelman
880	Navel gazing into the future	2012/4/16	Annalise Toberman
881	Pigs might fly	2012/3/30	Kate Anderson
882	MRS 2012 – Greater insight in sight	2012/3/22	Brian Kavanagh
883	If it's your second largest market can it really be called "emerging" anymore?	2012/3/19	Phil Stubington
884	Innovation and the commodity product – really?	2012/3/8	Sara Spinks
885	Time for a big rethink	2012/3/8	Kate Anderson
886	Fast track to improvement	2012/2/27	Charlotte Crichton
887	Freedom or Failure?	2012/2/17	Clayton Fussell
888	Innovation through iteration and triangulation	2012/2/15	Kate Downer
889	Is the Daily Mail smarter than it looks?	2012/2/10	Louise Amantani
890	Cloud: buzz word or dirty word?	2012/2/6	Martin Wootton
891	Our New Tool Makes It Easy to Determine Your LTL Freight Class	2012/6/7	Logan Andres
892	Survival story of FedEx courier's daughter provides extra motivation for volunteer work with 2012 Jimmy Fund Scooper Bowl	2012/6/6	Null
893	Facility Expansion and Infrastructure Improvements Continue at MEM World HUB	2012/6/4	John Dunavant
894	"E-tail" Likely to Overtake Traditional Retail	2012/5/30	Gene Huang
895	Our Office is Your Office: FedEx Office Reaches Out to Small Business Community with Tweet Chats	2012/5/23	David Kassen
896	The Three C's of Success: Culture, Community, Caring	2012/5/22	Bill Lovell
897	We May Be Big, But We Think Small…	2012/5/21	Denise Yunkun
898	Uncovering the Heart of FedEx	2012/5/21	Lisa Lisson
899	EarthSmart Innovations	2012/5/18	Mitch Jackson
900	"What's Your Sign?" – Not Just a Conversation Opener	2012/5/14	Kellie Graddy
901	Connect You: TechWomen at LinkedIn	2012/9/21	Florina Xhabija
902	The Best Way to Network with Alumni on LinkedIn	2012/9/20	Lindsey Pollak
903	Why LinkedIn Joined The Internet Association	2012/9/20	Eric Heath
904	LinkedIn Board Connect: Helping Nonprofits Find Board Members on LinkedIn	2012/9/17	Meg Garlinghouse

905	Hey UK, You're One in 10 Million!	2012/9/16	Ariel Eckstein
906	Winning More Business on LinkedIn	2012/9/14	Sourov De
907	It's a Small Business World After All: How to Make More Money on LinkedIn	2012/9/13	Nicole Williams
908	Introducing a New Look for Company Pages	2012/9/6	Mike Grishaver
909	More LinkedIn Mobile Goodness	2012/9/5	Joff Redfern
910	New: Get Notified on LinkedIn	2012/9/5	Angela Yang
911	Flying Across the World for a Summer Internship	2012/8/31	Yuan Daisy Gao
912	The Great LinkedIn Job Swap	2012/8/30	Rich Wong
913	Innovating with the LinkedIn Idea Bank	2012/8/29	Stefan Krawczyk
914	Going Global with LinkedIn Ads	2012/8/28	Gyanda Sachdeva
915	An Intern's Recap of LinkedIn Intern Hackday	2012/8/27	Karan Parikh
916	6 Steps to Shipping a Product to Your Most Important Customers	2012/8/24	Vicki Slavina
917	How to Network with VIPs on LinkedIn	2012/8/23	Lindsey Pollak
918	New Adventures with LinkedIn: A Retiree's Journey	2012/8/22	Howard Silvers
919	Living the Dream as a LinkedIn Intern	2012/8/17	Gilbert Hernandez
920	Taking My Online Professional Network Offline	2012/8/15	Linda Descano
921	Building a Digital Classroom on LinkedIn	2012/8/14	Leonardo Brant
922	5 Things I Learned as a Data Science Intern	2012/8/10	Nihit Desai
923	First Women's Hackday at LinkedIn	2012/8/10	Florina Xhabija
924	LinkedIn Joins the New Office App Store	2012/8/9	David Breger
925	MBA Life at LinkedIn	2012/8/8	Lucy Zang
926	LinkedIn Platform Further Enables Professional Content Sharing	2012/8/7	Madhu Gupta
927	Employing America, Finding Unity	2012/8/6	Daniel Roth
928	LinkedIn's Q2 2012 Earnings Call	2012/8/2	Steve Sordello
929	Stereotypes	2012/9/20	Jeffrey Wood
930	Snaps Image - Part 2	2012/8/27	Ed Nicholson
931	Tyson Donates 193,000 Ponds of Protein to Nationwide Food Banks	2012/8/23	Brady Tackett
932	Monsanto, ADM And The Fight Against Hunger. A Good Partnership?	2012/8/16	Ed Nicholson
933	How Will We Feed More People With Fewer Resources?	2012/8/14	Ed Nicholson
934	Does SNAP Have An Image Problem?	2012/8/6	Ed Nicholson
935	Real Hunger People Don't Talk.	2012/7/5	Ed Nicholson
936	Where's The Online Discussion About Hunger?	2012/6/29	Ed Nicholson
937	Partnering In Mississippi	2012/6/25	Ed Nicholson
938	Fewer Kids Hungry In Arkansas	2012/6/14	Ed Nicholson
939	Return To Joplin	2012/5/31	Ed Nicholson
940	In Memoriam: Sue Brockway	2012/5/3	Ed Nicholson

941	Chef Aaron McCargo, Jr. KNOW Hunger Video	2012/4/30	Ed Nicholson
942	You Shared Your KNOWledge–Austin Got Food	2012/4/25	Danielle Smith
943	You Shared Your KNOWledge–West Michigan Got Food	2012/4/25	Danielle Smith
944	You Shared Your KNOWledge–Houston Got Food	2012/4/25	Danielle Smith
945	You Shared Your KNOWledge–Southeastern Michigan Got Food	2012/4/25	Danielle Smith
946	You Shared Your KNOWledge–LA Got Food	2012/4/24	Ed Nicholson
947	Share This Message And Help Feed Hungry People In Los Angeles.	2012/4/18	Danielle Smith
948	Share This Message And Help Feed Hungry People In Austin.	2012/4/18	Danielle Smith
949	Share This Message And Help Feed Hungry People In New York City.	2012/4/18	Danielle Smith
950	Share This Message And Help Feed Hungry People In West Michigan.	2012/4/18	Danielle Smith
951	Share This Message And Help Feed Hungry People In Orlando.	2012/4/18	Danielle Smith
952	Share This Message And Help Feed Hungry People In New York.	2012/4/18	Danielle Smith
953	Share This Message And Help Feed Hungry People In Houston.	2012/4/18	Danielle Smith
954	Toasty Employees = Productive Employees	2012/9/21	Julia Ehrenfeld
955	Google's Picture Perfect Acquisition	2012/9/19	Aaron Steinfeld
956	Media Navel Gaze: September 17, 2012	2012/9/17	Mark Kollar
957	Nothing Lasts Forever:USA Today Boldly Overhauls Brand	2012/9/14	Sean Silva
958	20 Years Ago This Week:My Ride Across America	2012/9/11	Russell Sherman
959	Media Navel Gaze: September 10, 2012	2012/9/10	Mark Kollar
960	The Opportunity Cost of Tuning Out	2012/9/5	Emily Sackett
961	Media Navel Gaze: September 4, 2012	2012/9/4	Mark Kollar
962	More Work, Less Time	2012/8/31	Olivia Offner
963	Graphic Information: Nintendo vs. Sega	2012/8/27	Aaron Steinfeld
964	Media Navel Gaze: August 27, 2012	2012/8/27	Mark Kollar
965	R U OK?	2012/8/24	Mark Kollar
966	Men in the World of PR	2012/8/24	Spencer Gotschall
967	Passing the Wall-to-Wall Coverage Baton…	2012/8/22	Nicole Bliman
968	What's #Trending Now: Olympic Texts	2012/8/21	Vu Chung
969	Media Navel Gaze: August 20, 2012	2012/8/20	Mark Kollar
970	The Importance of Being Curious	2012/8/17	Jake Daubenspeck

971	Unboxed Lunch: “Truckie” is the New Foodie	2012/8/16	Emily Sackett
972	Helen Gurley Brown:An Unexpected Inspiration	2012/8/15	Jennifer Prosek
973	Media Navel Gaze: August 13, 2012	2012/8/13	Mark Kollar
974	Italicizing Caution	2012/8/8	Olivia Offner
975	Media Navel Gaze: August 6, 2012	2012/8/6	Mark Kollar
976	In Fantasy Football and Finance, Remember to Tune Out The Noise	2012/8/3	Joshua Passman
977	Geocaching & PR: Not So Different After All	2012/8/2	Jamie Kloss
978	The Flip Side	2012/8/1	Olivia Offner
979	MobiMOOC – Join The Corporate m-Learning Week (24th – 28th Sep 2012)	2012/9/20	Amit Garg
980	4 Key Emerging Trends in LMS	2012/9/18	Amit Gautam
981	Is MOOC Suitable For Corporate & Workplace Learning?	2012/9/17	Amit Garg
982	Embedded Ubiquitous Learning	2012/9/13	Abhijit Kadle
983	The Mobile Learning Ecosystem I – 10 Important Questions	2012/9/5	Abhijit Kadle
984	7 Key Components Of The LMS Switching Process	2012/9/4	Amit Gautam
985	3 Reasons To Switch Your LMS	2012/8/29	Jason Clymer
986	Free Mobile Learning Course – MobiMOOC2012	2012/8/28	Amit Garg
987	Key Decision Areas for Implementing Mobile Learning – LearnX 2012 Session	2012/8/23	Amit Garg
988	No Escaping Mobile Learning – Webinar Recording & Audience Queries Answered	2012/8/22	Amit Garg
989	6 Steps To A Better Mobile Learning Strategy	2012/8/16	Amit Garg
990	Webinar Video – A Unified Approach To Learning Management	2012/8/7	Aneesh Bhat
991	You Cannot Escape Mobile Learning!	2012/8/2	Pranjalee Thanekar
992	Augmented Reality – Making Paper Interactive	2012/7/31	Abhijit Kadle
993	A Time Saving Assessment	2012/7/26	Mandar Joshi
994	ASTD Webcast On New Mobile Learning Research (Watch Recording)	2012/7/24	Amit Garg
995	10 Effective LMS RFP Guidelines	2012/7/19	Dhaval Trivedi
996	7 Reasons Why You Need To Adopt A Unified Learning System	2012/7/17	Pranjalee Thanekar
997	10 Great Learning Game Design Links	2012/7/11	Abhijit Kadle1
998	LMS – Evolving Into A Unified Learning System	2012/7/10	Pranjalee Thanekar
999	Top 7 Myths Of Mobile Learning	2012/7/5	Amit Garg
1000	Is It Time To Move Onto A Unified Learning System?	2012/7/4	Aneesh Bhat

1001	Leveraging Mobile Learning For Sales Training	2012/7/3	Aneesh Bhat
1002	mLearnCon 2012	2012/7/2	Abhijit Kadle
1003	More Mobile Numbers	2012/6/28	Abhijit Kadle
1004	New Learning Technology Research – Be A Part Of It	2012/6/25	Amit Garg
1005	Tin Can: My First Impressions From mLearnCon 2012	2012/6/21	Amit Gautam
1006	Microsoft Surface: Will It Help mLearning?	2012/6/20	Amit Garg
1007	mLearnCon 2011 Recap	2012/6/14	Abhijit Kadle
1008	5 Tips For Managing Effective Blended Learning Through An LMS	2012/6/13	Amol Shinde
1009	Is Paper History	2012/9/18	Link Hoewing
1010	Some Thoughts About IP Interconnection – Blog Post Number One	2012/9/4	Link Hoewing
1011	If you build it....	2012/6/21	Link Hoewing
1012	Verizon and AT&T statement regarding ARCEP's attempt to regulate Internet peering and transit agreements	2012/6/21	Link Hoewing
1013	A Visit to Verizon’s Ashburn Campus	2012/6/13	Link Hoewing
1014	Things That Don’t Stick	2012/5/22	Link Hoewing
1015	Verizon’s Business Model and Shared Success	2012/5/1	Link Hoewing
1016	Prepared Remarks of Verizon EVP Tom Tauke to the NDN	2012/4/20	Link Hoewing
1017	Verizon Endorses Collaborative Effort to Address Theft of Mobile Devices and Protect Consumers	2012/4/10	John 'CZ' Czwartacki
1018	The New FTC Report on Privacy - Our View	2012/4/10	Link Hoewing
1019	Today’s Three Screen Minimum	2012/3/30 10:34	John 'CZ' Czwartacki
1020	SXSW Interactive: A Microcosm of Today’s Wired World	2012/3/30	John 'CZ' Czwartacki
1021	Crowe Says No Spectrum Shortage/ Really?	2012/3/29	David Young
1022	Political Silly Season	2012/3/29	Link Hoewing
1023	Getting Unused Spectrum in the Hands of Customers	2012/3/22	John 'CZ' Czwartacki
1024	The iPad’s Journey to Today	2012/3/14	Link Hoewing
1025	Interconnecting with the Future	2012/2/24	David Young
1026	Spectrum - Crunching the Numbers	2012/2/23	Charla Rath
1027	What a Difference a Decade Makes	2012/2/23	Link Hoewing
1028	Baby Steps: Overcoming State Licensure Obstacles to Telemedicine	2012/2/17	John 'CZ' Czwartacki
1029	Spectrum Provisions Bring Consumer Benefits, New Investments, Much-Needed Public Safety Tools	2012/2/16	John 'CZ' Czwartacki

1030	Chickens and Making Progress in Solving Tough Policy Issues	2012/2/15	Link Hoewing
1031	Have You Hugged a Scientist Today?	2012/2/10	John 'CZ' Czwartacki
1032	The *Perfect Storm* That Led to the Explosion of Apps	2012/2/10	Link Hoewing
1033	Washington CAN Get it When it Comes to the Internet - Here's How	2012/1/31	Link Hoewing
1034	Can Shopping for Healthcare be as Easy as Planning a Trip Online?	2012/9/24	Jeff Hechemy
1035	High-Volume Production Print with the Xerox iGen™ 150 Press	2012/9/21	Null
1036	What Would NASA Use? – A paper that can take it!	2012/9/20	Liam Cummings
1037	Connecting With Customers, Part 3	2012/9/20	James Dunn
1038	Four Hidden Productivity Killers (and How To Fix Them)	2012/9/19	Doryen Edward Chin
1039	Cracking the Marketer's Code	2012/9/19	Mary Roddy
1040	How to Properly Prepare for a Meet-and-Confer Session	2012/9/18	Null
1041	You Don't Have to Become a Marketing Services Provider to Offer Marketing Services	2012/9/18	Howard Fenton
1042	Xerox Is Mobile: Xerox Mobile Print Solution	2012/9/18	Null
1043	Power of Data	2012/9/17	Null
1044	Graph Expo 2012	2012/9/17	Null
1045	Xerox Device Agent Lite	2012/9/14	Null
1046	Building a Team of Virtual Workers: 4 Issues to Avoid	2012/9/12	Matt Keener
1047	Kodak makes news with On Demand Books – and that's good news for Xerox	2012/9/12	John Conley
1048	A One-Man-Band Inside Your Multifunction Printer	2012/9/12	Paul Criswell
1049	Building a Team of Virtual Workers: 4 Issues to Avoid	2012/9/12	Matt Keener
1050	Is Digital Printing Killing Offset Printing?	2012/9/11	Howard Fenton
1051	Print Your Book On Demand – Espresso Book Machine a Xerox Solution	2012/9/11	Null
1052	Banking in Latin America	2012/9/10	Null
1053	Five Ways to Improve the Efficiency of E-Discovery	2012/9/10	Null
1054	Learning From Your Company's "Collective Genius" – It Starts with a Question -1	2012/9/7	Breanna Banford
1055	Flashback Friday: The Xerox Copyflo 11	2012/9/7	Christina Vullo

1056	Learning From Your Company's "Collective Genius" – It Starts with a Question -2	2012/9/7	Breanna Banford
1057	Yes, Xerox Has Online Support	2012/9/7	Null
1058	Why Keyword Search Won't Go Away	2012/9/5	Null

Appendix 2: Linguistic features selected for the MD analysis of corporate blogs

A. Tense and aspect markers
1 Past tense
2 Perfect aspect
3 Present tense
B. Place and time adverbials
4 Place adverbials (e.g., above, beside, outdoors)
5 Time adverbials (e.g., early, instantly, soon)
C. Pronouns and pro-verbs
6 First-person pronouns
7 Second-person pronouns
8 Third-person personal pronouns (excluding it)
9 Pronoun it
10 Demonstrative pronouns (that, this, these, those as pronouns)
11 Indefinite pronouns (e.g., anybody, nothing, someone)
D. Questions
12 Direct WH questions
E. Nominal forms
13 Total nouns
F. Passives
14. Agentless passives
15. By-passives
G. Stative forms
16. Existential there
H. Subordination features
17. Infinitves
18. THAT relative clauses controlled by a verb
19. THAT relative clauses controlled by an adjective
20. THAT relative clauses controlled by a verb
21. WH relative clauses
22. Concessive adverbial subordinators (although, though)
23. Causative adverbial subordinators (because)
24. Conditional adverbial subordinators (if, unless)
J. Prepositional phrases, adjectives, and adverbs
25. Total prepositional phrases
26. Total adjectives
27. Total adverbs
K. Lexical specificity
28. Type-token ratio
29. Mean word length

L. Modals

30. Possibility modals (can, may, might, could)

31 Necessity modals (ought, should, must)

32. Predictive modals (will, would, shall)

Appendix 3: A broad structure of International Standard Industrial Classification (ISIC) of All Economic Activities

Sections	Divisions	Description
A	01–03	Agriculture, forestry and fishing
B	05–09	Mining and quarrying
C	10–33	Manufacturing
D	35	Electricity, gas, steam and air conditioning supply
E	36–39	Water supply; sewerage, waste management and remediation activities
F	41–43	Construction
G	45–47	Wholesale and retail trade; repair of motor vehicles and motorcycles
H	49–53	Transportation and storage
I	55–56	Accommodation and food service activities
J	58–63	Information and communication
K	64–66	Financial and insurance activities
L	68	Real estate activities
M	69–75	Professional, scientific and technical activities
N	77–82	Administrative and support service activities
O	84	Public administration and defence; compulsory social security
P	85	Education
Q	86–88	Human health and social work activities
R	90–93	Arts, entertainment and recreation
S	94–96	Other service activities
T	97–98	Activities of households as employers; undifferentiated goods- and services-producing activities of households for own use
U	99	Activities of extraterritorial organizations and bodies

Appendix 4: Global Industry Classification Standard (GICS)

Code	Sector	Subcode	Industry Groups
10	Energy	1010	Energy
15	Materials	1510	Materials
20	Industrials	2010	Capital Goods
		2020	Commercial & Professional Services
		2030	Transportation
25	Consumer Discretionary	2510	Automobiles & Components
		2520	Consumer Durables & Apparel
		2530	Hotels Restaurants & Leisure
		2540	Media
		2550	Retailing
30	Consumer Staples	3010	Food & Staples Retailing
		3020	Food, Beverage & Tobacco

		3030	Household & Personal Products
35	Health Care	3510	Health Care Equipment & Services
		3520	Pharmaceuticals & Biotechnology
40	Financials	4010	Banks
		4020	Diversified Financials
		4030	Insurance
		4040	Real Estate
45	Information Technology	4510	Software & Services
		4520	Technology Hardware & Equipment
		4530	Semiconductors & Semiconductor Equipment
50	Telecommunication Services	5010	Telecommunication Services
55	Utilities	5510	Utilities

Appendix 5: A full list of CB texts categorized into clusters and industry groups

CB Text No.	The names of CB	Cluster (Text type)	Industry
1	Delta Air Lines Blog	5	Service industry
2	Delta Air Lines Blog	5	Service industry
3	Delta Air Lines Blog	1	Service industry
4	Delta Air Lines Blog	2	Service industry
5	Delta Air Lines Blog	5	Service industry
6	Delta Air Lines Blog	2	Service industry
7	Delta Air Lines Blog	6	Service industry
8	Delta Air Lines Blog	6	Service industry
9	Delta Air Lines Blog	6	Service industry
10	Delta Air Lines Blog	1	Service industry
11	Delta Air Lines Blog	6	Service industry
12	Delta Air Lines Blog	5	Service industry
13	Delta Air Lines Blog	6	Service industry
14	Delta Air Lines Blog	6	Service industry
15	Delta Air Lines Blog	1	Service industry
16	Delta Air Lines Blog	5	Service industry
17	Delta Air Lines Blog	3	Service industry
18	Delta Air Lines Blog	3	Service industry
19	Delta Air Lines Blog	2	Service industry
20	Delta Air Lines Blog	6	Service industry
21	Whole Story	6	Wholesale and retail trade
22	Whole Story	6	Wholesale and retail trade
23	Whole Story	6	Wholesale and retail trade
24	Whole Story	3	Wholesale and retail trade
25	Whole Story	6	Wholesale and retail trade
26	Whole Story	2	Wholesale and retail trade
27	Whole Story	6	Wholesale and retail trade
28	Whole Story	6	Wholesale and retail trade
29	Whole Story	1	Wholesale and retail trade
30	Whole Story	6	Wholesale and retail trade
31	Whole Story	2	Wholesale and retail trade
32	Whole Story	6	Wholesale and retail trade
33	Whole Story	3	Wholesale and retail trade
34	Whole Story	4	Wholesale and retail trade
35	Whole Story	6	Wholesale and retail trade

36	Whole Story	1	Wholesale and retail trade
37	Whole Story	2	Wholesale and retail trade
38	Whole Story	6	Wholesale and retail trade
39	Whole Story	5	Wholesale and retail trade
40	Whole Story	6	Wholesale and retail trade
41	Marriott On the Move	2	Service industry
42	Marriott On the Move	2	Service industry
43	Marriott On the Move	5	Service industry
44	Marriott On the Move	2	Service industry
45	Marriott On the Move	1	Service industry
46	Marriott On the Move	2	Service industry
47	Marriott On the Move	1	Service industry
48	Marriott On the Move	2	Service industry
49	Marriott On the Move	2	Service industry
50	Marriott On the Move	2	Service industry
51	Marriott On the Move	2	Service industry
52	Marriott On the Move	2	Service industry
53	Marriott On the Move	1	Service industry
54	Marriott On the Move	2	Service industry
55	Marriott On the Move	1	Service industry
56	Marriott On the Move	2	Service industry
57	Marriott On the Move	2	Service industry
58	Marriott On the Move	1	Service industry
59	Marriott On the Move	2	Service industry
60	Marriott On the Move	2	Service industry
61	The Cleanest Line	2	Manufacturing
62	The Cleanest Line	1	Manufacturing
63	The Cleanest Line	2	Manufacturing
64	The Cleanest Line	2	Manufacturing
65	The Cleanest Line	2	Manufacturing
66	The Cleanest Line	.	
67	The Cleanest Line	1	Manufacturing
68	The Cleanest Line	5	Manufacturing
69	The Cleanest Line	2	Manufacturing
70	The Cleanest Line	2	Manufacturing
71	The Cleanest Line	2	Manufacturing
72	The Cleanest Line	2	Manufacturing
73	The Cleanest Line	4	Manufacturing
74	The Cleanest Line	2	Manufacturing
75	The Cleanest Line	2	Manufacturing
76	The Cleanest Line	1	Manufacturing
77	The Cleanest Line	2	Manufacturing
78	The Cleanest Line	5	Manufacturing

79	The Cleanest Line	5	Manufacturing
80	The Cleanest Line	2	Manufacturing
81	Toyota Open Road Blog	1	Manufacturing
82	Toyota Open Road Blog	.	
83	Toyota Open Road Blog	2	Manufacturing
84	Toyota Open Road Blog	1	Manufacturing
85	Toyota Open Road Blog	2	Manufacturing
86	Toyota Open Road Blog	2	Manufacturing
87	Toyota Open Road Blog	1	Manufacturing
88	Toyota Open Road Blog	2	Manufacturing
89	Toyota Open Road Blog	2	Manufacturing
90	Toyota Open Road Blog	2	Manufacturing
91	Red Cross Chat	1	Manufacturing
92	Red Cross Chat	.	
93	Red Cross Chat	.	
94	Red Cross Chat	2	Manufacturing
95	Red Cross Chat	1	Manufacturing
96	Red Cross Chat	2	Manufacturing
97	Red Cross Chat	1	Manufacturing
98	Red Cross Chat	1	Manufacturing
99	Red Cross Chat	1	Manufacturing
100	Red Cross Chat	2	Manufacturing
101	Red Cross Chat	2	Activities of extraterritorial organizations and bodies
102	Red Cross Chat	2	Activities of extraterritorial organizations and bodies
103	Red Cross Chat	1	Activities of extraterritorial organizations and bodies
104	Red Cross Chat	4	Activities of extraterritorial organizations and bodies
105	Red Cross Chat	1	Activities of extraterritorial organizations and bodies
106	Red Cross Chat	4	Activities of extraterritorial organizations and bodies
107	Red Cross Chat	2	Activities of extraterritorial organizations and bodies
108	Red Cross Chat	2	Activities of extraterritorial organizations and bodies
109	Red Cross Chat	5	Activities of extraterritorial organizations and bodies
110	Red Cross Chat	2	Activities of extraterritorial organizations and bodies
111	Red Cross Chat	1	Activities of extraterritorial

			organizations and bodies
112	Red Cross Chat	2	Activities of extraterritorial organizations and bodies
113	Red Cross Chat	4	Activities of extraterritorial organizations and bodies
114	Red Cross Chat	2	Activities of extraterritorial organizations and bodies
115	Red Cross Chat	2	Activities of extraterritorial organizations and bodies
116	Red Cross Chat	2	Activities of extraterritorial organizations and bodies
117	Red Cross Chat	5	Activities of extraterritorial organizations and bodies
118	Red Cross Chat	1	Activities of extraterritorial organizations and bodies
119	Red Cross Chat	2	Activities of extraterritorial organizations and bodies
120	Red Cross Chat	2	Activities of extraterritorial organizations and bodies
121	FedEx Citizenship Blog	1	Service industry
122	FedEx Citizenship Blog	1	Service industry
123	FedEx Citizenship Blog	2	Service industry
124	FedEx Citizenship Blog	2	Service industry
125	FedEx Citizenship Blog	2	Service industry
126	FedEx Citizenship Blog	2	Service industry
127	FedEx Citizenship Blog	1	Service industry
128	FedEx Citizenship Blog	4	Service industry
129	FedEx Citizenship Blog	2	Service industry
130	FedEx Citizenship Blog	1	Service industry
131	FedEx Citizenship Blog	2	Service industry
132	FedEx Citizenship Blog	2	Service industry
133	FedEx Citizenship Blog	1	Service industry
134	FedEx Citizenship Blog	1	Service industry
135	FedEx Citizenship Blog	2	Service industry
136	FedEx Citizenship Blog	2	Service industry
137	FedEx Citizenship Blog	1	Service industry
138	FedEx Citizenship Blog	2	Service industry
139	FedEx Citizenship Blog	2	Service industry
140	FedEx Citizenship Blog	2	Service industry
141	Facebook Blog	5	E-commerce and online social networking service
142	Facebook Blog	5	E-commerce and online social networking service

143	Facebook Blog	5	E-commerce and online social networking service
144	Facebook Blog	5	E-commerce and online social networking service
145	Facebook Blog	5	E-commerce and online social networking service
146	Facebook Blog	1	E-commerce and online social networking service
147	Facebook Blog	2	E-commerce and online social networking service
148	Facebook Blog	5	E-commerce and online social networking service
149	Facebook Blog	5	E-commerce and online social networking service
150	Facebook Blog	5	E-commerce and online social networking service
151	Facebook Blog	5	E-commerce and online social networking service
152	Facebook Blog	5	E-commerce and online social networking service
153	Facebook Blog	5	E-commerce and online social networking service
154	Facebook Blog	5	E-commerce and online social networking service
155	Facebook Blog	5	E-commerce and online social networking service
156	Facebook Blog	5	E-commerce and online social networking service
157	Facebook Blog	5	E-commerce and online social networking service
158	Facebook Blog	5	E-commerce and online social networking service
159	Facebook Blog	5	E-commerce and online social networking service
160	Facebook Blog	5	E-commerce and online social networking service
161	Turkey Hill	5	Wholesale and retail trade
162	Turkey Hill	1	Wholesale and retail trade
163	Turkey Hill	5	Wholesale and retail trade
164	Turkey Hill	2	Wholesale and retail trade
165	Turkey Hill	2	Wholesale and retail trade
166	Turkey Hill	2	Wholesale and retail trade
167	Turkey Hill	2	Wholesale and retail trade

168	Turkey Hill	1	Wholesale and retail trade
169	Turkey Hill	5	Wholesale and retail trade
170	Turkey Hill	2	Wholesale and retail trade
171	Turkey Hill	4	Wholesale and retail trade
172	Turkey Hill	5	Wholesale and retail trade
173	Turkey Hill	2	Wholesale and retail trade
174	Turkey Hill	2	Wholesale and retail trade
175	Turkey Hill	2	Wholesale and retail trade
176	Turkey Hill	1	Wholesale and retail trade
177	Turkey Hill	4	Wholesale and retail trade
178	Turkey Hill	2	Wholesale and retail trade
179	Turkey Hill	1	Wholesale and retail trade
180	Turkey Hill	5	Wholesale and retail trade
181	Lede	1	Publishing industry
182	Lede	1	Publishing industry
183	Lede	4	Publishing industry
184	Lede	4	Publishing industry
185	Lede	2	Publishing industry
186	Wallet	4	Publishing industry
187	Wallet	1	Publishing industry
188	Wallet	2	Publishing industry
189	Wallet	4	Publishing industry
190	Comment is free | guardian.co.uk	5	Publishing industry
191	Comment is free | guardian.co.uk	.	
192	Comment is free | guardian.co.uk	4	Publishing industry
193	Comment is free | guardian.co.uk	4	Publishing industry
194	Comment is free | guardian.co.uk	1	Publishing industry
195	Comment is free | guardian.co.uk	5	Publishing industry
196	Lede	1	Publishing industry
197	Lede	1	Publishing industry
198	Lede	1	Publishing industry
199	Lede	2	Publishing industry
200	Lede	1	Publishing industry
201	Delta Air Lines Blog	2	Service industry
202	Delta Air Lines Blog	2	Service industry
203	Delta Air Lines Blog	1	Service industry
204	Delta Air Lines Blog	2	Service industry

205	Delta Air Lines Blog	4	Service industry
206	Whole Story	2	Wholesale and retail trade
207	Whole Story	2	Wholesale and retail trade
208	Whole Story	4	Wholesale and retail trade
209	Whole Story	2	Wholesale and retail trade
210	Whole Story	2	Wholesale and retail trade
211	Amazon Web Services Blog	2	E-commerce and online social networking service
212	Amazon Web Services Blog	5	E-commerce and online social networking service
213	Amazon Web Services Blog	2	E-commerce and online social networking service
214	Amazon Web Services Blog	4	E-commerce and online social networking service
215	Amazon Web Services Blog	4	E-commerce and online social networking service
216	Amazon Web Services Blog	4	E-commerce and online social networking service
217	Amazon Web Services Blog	2	E-commerce and online social networking service
218	Amazon Web Services Blog	5	E-commerce and online social networking service
219	Amazon Web Services Blog	2	E-commerce and online social networking service
220	Amazon Web Services Blog	2	E-commerce and online social networking service
221	Amazon Web Services Blog	4	E-commerce and online social networking service
222	Amazon Web Services Blog	2	E-commerce and online social networking service
223	Amazon Web Services Blog	2	E-commerce and online social networking service
224	Amazon Web Services Blog	5	E-commerce and online social networking service
225	Amazon Web Services Blog	2	E-commerce and online social networking service
226	Amazon Web Services Blog	2	E-commerce and online social networking service
227	Amazon Web Services Blog	2	E-commerce and online social networking service
228	Amazon Web Services Blog	5	E-commerce and online social networking service
229	Amazon Web Services Blog	2	E-commerce and online social

			networking service
230	Amazon Web Services Blog	2	E-commerce and online social networking service
231	Amazon Web Services Blog	2	E-commerce and online social networking service
232	Amazon Web Services Blog	2	E-commerce and online social networking service
233	Amazon Web Services Blog	2	E-commerce and online social networking service
234	Amazon Web Services Blog	5	E-commerce and online social networking service
235	Amazon Web Services Blog	2	E-commerce and online social networking service
236	Avaya - The Blog	2	Information Technology
237	Avaya - The Blog	5	Information Technology
238	Avaya - The Blog	4	Information Technology
239	Avaya - The Blog	4	Information Technology
240	Avaya - The Blog	5	Information Technology
241	Avaya - The Blog	4	Information Technology
242	Avaya - The Blog	4	Information Technology
243	Avaya - The Blog	4	Information Technology
244	Avaya - The Blog	2	Information Technology
245	Avaya - The Blog	4	Information Technology
246	Avaya - The Blog	2	Information Technology
247	Avaya - The Blog	4	Information Technology
248	Avaya - The Blog	2	Information Technology
249	Avaya - The Blog	1	Information Technology
250	Avaya - The Blog	2	Information Technology
251	Avaya - The Blog	2	Information Technology
252	Avaya - The Blog	2	Information Technology
253	Avaya - The Blog	4	Information Technology
254	Avaya - The Blog	2	Information Technology
255	Avaya - The Blog	4	Information Technology
256	Avaya - The Blog	5	Information Technology
257	Avaya - The Blog	2	Information Technology
258	Avaya - The Blog	4	Information Technology
259	Avaya - The Blog	2	Information Technology
260	Avaya - The Blog	4	Information Technology
261	Brand Channeler	2	Information and communication
262	Brand Channeler	5	Information and communication
263	Brand Channeler	4	Information and communication
264	Brand Channeler	2	Information and communication
265	Brand Channeler	2	Information and communication

266	Brand Channeler	4	Information and communication
267	Brand Channeler	2	Information and communication
268	Brand Channeler	4	Information and communication
269	Brand Channeler	2	Information and communication
270	Brand Channeler	5	Information and communication
271	Brand Channeler	2	Information and communication
272	Brand Channeler	2	Information and communication
273	Brand Channeler	5	Information and communication
274	Brand Channeler	5	Information and communication
275	Brand Channeler	2	Information and communication
276	Brand Channeler	4	Information and communication
277	Brand Channeler	2	Information and communication
278	Brand Channeler	2	Information and communication
279	Brand Channeler	2	Information and communication
280	Brand Channeler	4	Information and communication
281	Brand Channeler	2	Information and communication
282	Brand Channeler	2	Information and communication
283	Brand Channeler	2	Information and communication
284	Brand Channeler	4	Information and communication
285	Brand Channeler	2	Information and communication
286	BreakingPoint Labs Blog	5	Information Technology
287	BreakingPoint Labs Blog	2	Information Technology
288	BreakingPoint Labs Blog	2	Information Technology
289	BreakingPoint Labs Blog	2	Information Technology
290	BreakingPoint Labs Blog	2	Information Technology
291	BreakingPoint Labs Blog	5	Information Technology
292	BreakingPoint Labs Blog	2	Information Technology
293	BreakingPoint Labs Blog	2	Information Technology
294	BreakingPoint Labs Blog	2	Information Technology
295	BreakingPoint Labs Blog	2	Information Technology
296	BreakingPoint Labs Blog	2	Information Technology
297	BreakingPoint Labs Blog	4	Information Technology
298	BreakingPoint Labs Blog	4	Information Technology
299	BreakingPoint Labs Blog	2	Information Technology
300	BreakingPoint Labs Blog	2	Information Technology
301	BreakingPoint Labs Blog	4	Information Technology
302	BreakingPoint Labs Blog	4	Information Technology
303	BreakingPoint Labs Blog	2	Information Technology
304	BreakingPoint Labs Blog	5	Information Technology
305	BreakingPoint Labs Blog	4	Information Technology
306	BreakingPoint Labs Blog	2	Information Technology
307	BreakingPoint Labs Blog	4	Information Technology
308	BreakingPoint Labs Blog	2	Information Technology

309	BreakingPoint Labs Blog	2	Information Technology
310	BreakingPoint Labs Blog	2	Information Technology
311	Business Property	5	Service industry
312	Business Property	2	Service industry
313	Business Property	4	Service industry
314	Business Property	5	Service industry
315	Business Property	4	Service industry
316	Business Property	4	Service industry
317	Business Property	5	Service industry
318	Business Property	2	Service industry
319	Business Property	2	Service industry
320	Business Property	5	Service industry
321	Business Property	5	Service industry
322	Business Property	4	Service industry
323	Business Property	4	Service industry
324	Business Property	2	Service industry
325	Business Property	4	Service industry
326	Business Property	2	Service industry
327	Business Property	5	Service industry
328	Business Property	5	Service industry
329	Business Property	2	Service industry
330	Business Property	5	Service industry
331	Business Property	2	Service industry
332	Business Property	5	Service industry
333	Business Property	4	Service industry
334	Business Property	5	Service industry
335	Business Property	4	Service industry
336	Clorox-Dr. Laundry	5	Manufacturing
337	Clorox-Dr. Laundry	5	Manufacturing
338	Clorox-Dr. Laundry	5	Manufacturing
339	Clorox-Dr. Laundry	2	Manufacturing
340	Clorox-Dr. Laundry	5	Manufacturing
341	Clorox-Dr. Laundry	5	Manufacturing
342	Clorox-Dr. Laundry	5	Manufacturing
343	Clorox-Dr. Laundry	5	Manufacturing
344	Clorox-Dr. Laundry	5	Manufacturing
345	Clorox-Dr. Laundry	5	Manufacturing
346	Clorox-Dr. Laundry	5	Manufacturing
347	Clorox-Dr. Laundry	5	Manufacturing
348	Clorox-Dr. Laundry	5	Manufacturing
349	Clorox-Dr. Laundry	5	Manufacturing
350	Clorox-Dr. Laundry	5	Manufacturing
351	Clorox-Dr. Laundry	5	Manufacturing

352	Clorox-Dr. Laundry	5	Manufacturing
353	Clorox-Dr. Laundry	5	Manufacturing
354	Clorox-Dr. Laundry	5	Manufacturing
355	Clorox-Dr. Laundry	2	Manufacturing
356	Clorox-Dr. Laundry	5	Manufacturing
357	Clorox-Dr. Laundry	5	Manufacturing
358	Clorox-Dr. Laundry	5	Manufacturing
359	Clorox-Dr. Laundry	5	Manufacturing
360	Clorox-Dr. Laundry	2	Manufacturing
361	ComScore Voice	2	Information and communication
362	ComScore Voice	2	Information and communication
363	ComScore Voice	4	Information and communication
364	ComScore Voice	2	Information and communication
365	ComScore Voice	2	Information and communication
366	ComScore Voice	4	Information and communication
367	ComScore Voice	4	Information and communication
368	ComScore Voice	4	Information and communication
369	ComScore Voice	1	Information and communication
370	ComScore Voice	4	Information and communication
371	ComScore Voice	2	Information and communication
372	ComScore Voice	4	Information and communication
373	ComScore Voice	4	Information and communication
374	ComScore Voice	4	Information and communication
375	ComScore Voice	4	Information and communication
376	ComScore Voice	4	Information and communication
377	ComScore Voice	4	Information and communication
378	ComScore Voice	2	Information and communication
379	ComScore Voice	1	Information and communication
380	ComScore Voice	2	Information and communication
381	ComScore Voice	4	Information and communication
382	ComScore Voice	2	Information and communication
383	ComScore Voice	2	Information and communication
384	Converge-A Bader Rutter Blog	2	Information and communication
385	Converge-A Bader Rutter Blog	1	Information and communication
386	Converge-A Bader Rutter Blog	5	Information and communication
387	Converge-A Bader Rutter Blog	5	Information and communication
388	Converge-A Bader Rutter Blog	5	Information and communication
389	Converge-A Bader Rutter Blog	2	Information and communication
390	Converge-A Bader Rutter Blog	2	Information and communication
391	Converge-A Bader Rutter Blog	4	Information and communication
392	Converge-A Bader Rutter Blog	2	Information and communication
393	Converge-A Bader Rutter Blog	2	Information and communication
394	Converge-A Bader Rutter Blog	4	Information and communication

395	Converge-A Bader Rutter Blog	2	Information and communication
396	Converge-A Bader Rutter Blog	2	Information and communication
397	Converge-A Bader Rutter Blog	2	Information and communication
398	Converge-A Bader Rutter Blog	4	Information and communication
399	Converge-A Bader Rutter Blog	2	Information and communication
400	Converge-A Bader Rutter Blog	1	Information and communication
401	Converge-A Bader Rutter Blog	.	
402	Converge-A Bader Rutter Blog	5	Information and communication
403	Converge-A Bader Rutter Blog	4	Information and communication
404	Converge-A Bader Rutter Blog	5	Information and communication
405	Converge-A Bader Rutter Blog	4	Information and communication
406	Converge-A Bader Rutter Blog	4	Information and communication
407	Converge-A Bader Rutter Blog	2	Information and communication
408	Corporate eye	2	Information and communication
409	Corporate eye	4	Information and communication
410	Corporate eye	1	Information and communication
411	Corporate eye	4	Information and communication
412	Corporate eye	2	Information and communication
413	Corporate eye	4	Information and communication
414	Corporate eye	1	Information and communication
415	Corporate eye	5	Information and communication
416	Corporate eye	4	Information and communication
417	Corporate eye	5	Information and communication
418	Corporate eye	1	Information and communication
419	Corporate eye	1	Information and communication
420	Corporate eye	5	Information and communication
421	Corporate eye	1	Information and communication
422	Corporate eye	2	Information and communication
423	Corporate eye	2	Information and communication
424	Corporate eye	1	Information and communication
425	Corporate eye	4	Information and communication
426	Corporate eye	1	Information and communication
427	Corporate eye	1	Information and communication
428	Corporate eye	4	Information and communication
429	Corporate eye	4	Information and communication
430	Dr. Laskys Blog	4	Manufacturing
431	Dr. Laskys Blog	1	Manufacturing
432	Dr. Laskys Blog	1	Manufacturing
433	Dr. Laskys Blog	4	Manufacturing
434	Dr. Laskys Blog	1	Manufacturing
435	Dr. Laskys Blog	1	Manufacturing
436	Dr. Laskys Blog	.	
437	Dr. Laskys Blog	1	Manufacturing

438	Dr. Laskys Blog	2	Manufacturing
439	Dr. Laskys Blog	4	Manufacturing
440	Dr. Laskys Blog	4	Manufacturing
441	Dr. Laskys Blog	2	Manufacturing
442	Dr. Laskys Blog	2	Manufacturing
443	Dr. Laskys Blog	2	Manufacturing
444	Dr. Laskys Blog	4	Manufacturing
445	Dr. Laskys Blog	4	Manufacturing
446	Dr. Laskys Blog	4	Manufacturing
447	Dr. Laskys Blog	4	Manufacturing
448	Dr. Laskys Blog	4	Manufacturing
449	Dr. Laskys Blog	1	Manufacturing
450	Fiskars Ambassador	4	Manufacturing
451	Fiskars Ambassador	.	
452	Fiskars Ambassador	5	Manufacturing
453	Fiskars Ambassador	4	Manufacturing
454	Fiskars Ambassador	4	Manufacturing
455	Fiskars Ambassador	5	Manufacturing
456	Fiskars Ambassador	.	
457	Fiskars Ambassador	5	Manufacturing
458	Fiskars Ambassador	5	Manufacturing
459	Fiskars Ambassador	5	Manufacturing
460	Fiskars Ambassador	5	Manufacturing
461	Fiskars Ambassador	5	Manufacturing
462	Fiskars Ambassador	2	Manufacturing
463	Fiskars Ambassador	5	Manufacturing
464	Fiskars Ambassador	5	Manufacturing
465	Fiskars Ambassador	2	Manufacturing
466	Fiskars Ambassador	2	Manufacturing
467	Fiskars Ambassador	2	Manufacturing
468	Fiskars Ambassador	2	Manufacturing
469	Fiskars Ambassador	2	Manufacturing
470	Fiskars Ambassador	1	Manufacturing
471	Fiskars Ambassador	5	Manufacturing
472	Fiskars Ambassador	5	Manufacturing
473	Fiskars Ambassador	5	Manufacturing
474	Fiskars Ambassador	2	Manufacturing
475	Fiskars Ambassador	5	Manufacturing
476	Fiskars Ambassador	2	Manufacturing
477	Fiskars Ambassador	5	Manufacturing
478	Fiskars Ambassador	4	Manufacturing
479	Fandango Movie Blog	1	E-commerce and online social networking service

480	Fandango Movie Blog	2	E-commerce and online social networking service
481	Fandango Movie Blog	1	E-commerce and online social networking service
482	Fandango Movie Blog	2	E-commerce and online social networking service
483	Fandango Movie Blog	2	E-commerce and online social networking service
484	Fandango Movie Blog	5	E-commerce and online social networking service
485	Fandango Movie Blog	4	E-commerce and online social networking service
486	Fandango Movie Blog	4	E-commerce and online social networking service
487	Fandango Movie Blog	.	
488	Fandango Movie Blog	4	E-commerce and online social networking service
489	Fandango Movie Blog	1	E-commerce and online social networking service
490	Fandango Movie Blog	2	E-commerce and online social networking service
491	Fandango Movie Blog	4	E-commerce and online social networking service
492	Fandango Movie Blog	1	E-commerce and online social networking service
493	Fandango Movie Blog	2	E-commerce and online social networking service
494	Fandango Movie Blog	2	E-commerce and online social networking service
495	Fandango Movie Blog	.	
496	Fandango Movie Blog	4	E-commerce and online social networking service
497	Fandango Movie Blog	5	E-commerce and online social networking service
498	Fandango Movie Blog	5	E-commerce and online social networking service
499	Fandango Movie Blog	5	E-commerce and online social networking service
500	Fandango Movie Blog	1	E-commerce and online social networking service
501	Fandango Movie Blog	4	E-commerce and online social networking service
502	Fandango Movie Blog	4	E-commerce and online social

			networking service
503	Fandango Movie Blog	4	E-commerce and online social networking service
504	Fandango Movie Blog	4	E-commerce and online social networking service
505	Fandango Movie Blog	5	E-commerce and online social networking service
506	Fandango Movie Blog	1	E-commerce and online social networking service
507	Fandango Movie Blog	1	E-commerce and online social networking service
508	Fandango Movie Blog	2	E-commerce and online social networking service
509	Turkey Hill	1	Wholesale and retail trade
510	Turkey Hill	.	
511	Turkey Hill	.	
512	Turkey Hill	5	Wholesale and retail trade
513	Turkey Hill	5	Wholesale and retail trade
514	Turkey Hill	.	
515	Turkey Hill	.	
516	Turkey Hill	.	
517	Turkey Hill	5	Wholesale and retail trade
518	Turkey Hill	2	Wholesale and retail trade
519	Turkey Hill	.	
520	Turkey Hill	4	Wholesale and retail trade
521	Turkey Hill	2	Wholesale and retail trade
522	Turkey Hill	2	Wholesale and retail trade
523	Turkey Hill	1	Wholesale and retail trade
524	Lede	1	Publishing industry
525	Lede	1	Publishing industry
526	Lede	1	Publishing industry
527	Lede	1	Publishing industry
528	Lede	4	Publishing industry
529	Lede	1	Publishing industry
530	Lede	2	Publishing industry
531	Lede	1	Publishing industry
532	Lede	1	Publishing industry
533	Lede	1	Publishing industry
534	Lede	1	Publishing industry
535	Lede	1	Publishing industry
536	Lede	1	Publishing industry
537	Garmin Blog	2	Manufacturing
538	Garmin Blog	2	Manufacturing

539	Garmin Blog	2	Manufacturing
540	Garmin Blog	2	Manufacturing
541	Garmin Blog	2	Manufacturing
542	Garmin Blog	4	Manufacturing
543	Garmin Blog	4	Manufacturing
544	Garmin Blog	5	Manufacturing
545	Garmin Blog	5	Manufacturing
546	Garmin Blog	4	Manufacturing
547	Garmin Blog	2	Manufacturing
548	Garmin Blog	.	
549	Garmin Blog	2	Manufacturing
550	Garmin Blog	2	Manufacturing
551	Garmin Blog	2	Manufacturing
552	Garmin Blog	.	
553	Garmin Blog	.	
554	Garmin Blog	5	Manufacturing
555	Garmin Blog	2	Manufacturing
556	Garmin Blog	2	Manufacturing
557	Garmin Blog	2	Manufacturing
558	Hu Yoshida	2	Manufacturing
559	Hu Yoshida	2	Manufacturing
560	Hu Yoshida	2	Manufacturing
561	Hu Yoshida	2	Manufacturing
562	Hu Yoshida	4	Manufacturing
563	Hu Yoshida	2	Manufacturing
564	Hu Yoshida	2	Manufacturing
565	Hu Yoshida	2	Manufacturing
566	Hu Yoshida	2	Manufacturing
567	Hu Yoshida	2	Manufacturing
568	Hu Yoshida	2	Manufacturing
569	Hu Yoshida	2	Manufacturing
570	Hu Yoshida	2	Manufacturing
571	Hu Yoshida	2	Manufacturing
572	Hu Yoshida	2	Manufacturing
573	Hu Yoshida	2	Manufacturing
574	Hu Yoshida	2	Manufacturing
575	Hu Yoshida	1	Manufacturing
576	Hu Yoshida	2	Manufacturing
577	Hu Yoshida	4	Manufacturing
578	IAGblog	2	Information and communication
579	IAGblog	2	Information and communication
580	IAGblog	5	Information and communication
581	IAGblog	2	Information and communication

582	IAGblog	4	Information and communication
583	IAGblog	2	Information and communication
584	IAGblog	2	Information and communication
585	IAGblog	2	Information and communication
586	IAGblog	2	Information and communication
587	IAGblog	2	Information and communication
588	IAGblog	5	Information and communication
589	IAGblog	1	Information and communication
590	IAGblog	2	Information and communication
591	IAGblog	2	Information and communication
592	IAGblog	2	Information and communication
593	IAGblog	2	Information and communication
594	IAGblog	2	Information and communication
595	IAGblog	4	Information and communication
596	IAGblog	2	Information and communication
597	IAGblog	5	Information and communication
598	IAGblog	2	Information and communication
599	IAGblog	1	Information and communication
600	IAGblog	4	Information and communication
601	IAGblog	.	
602	IAGblog	2	Information and communication
603	IAGblog	4	Information and communication
604	IAGblog	2	Information and communication
605	IAGblog	2	Information and communication
606	IAGblog	2	Information and communication
607	IAGblog	4	Information and communication
608	Infusionsoft Blog	5	Information and communication
609	Infusionsoft Blog	4	Information and communication
610	Infusionsoft Blog	2	Information and communication
611	Infusionsoft Blog	2	Information and communication
612	Infusionsoft Blog	4	Information and communication
613	Infusionsoft Blog	5	Information and communication
614	Infusionsoft Blog	2	Information and communication
615	Infusionsoft Blog	2	Information and communication
616	Infusionsoft Blog	5	Information and communication
617	Infusionsoft Blog	2	Information and communication
618	Infusionsoft Blog	5	Information and communication
619	Infusionsoft Blog	2	Information and communication
620	Infusionsoft Blog	4	Information and communication
621	Infusionsoft Blog	5	Information and communication
622	Infusionsoft Blog	5	Information and communication
623	Infusionsoft Blog	2	Information and communication
624	Infusionsoft Blog	2	Information and communication

625	Infusionsoft Blog	5	Information and communication
626	Infusionsoft Blog	5	Information and communication
627	Infusionsoft Blog	5	Information and communication
628	JNJ BTW	4	Service industry
629	JNJ BTW	4	Service industry
630	JNJ BTW	2	Service industry
631	JNJ BTW	2	Service industry
632	JNJ BTW	2	Service industry
633	JNJ BTW	5	Service industry
634	JNJ BTW	4	Service industry
635	JNJ BTW	2	Service industry
636	JNJ BTW	2	Service industry
637	JNJ BTW	2	Service industry
638	JNJ BTW	2	Service industry
639	JNJ BTW	2	Service industry
640	JNJ BTW	2	Service industry
641	JNJ BTW	4	Service industry
642	JNJ BTW	4	Service industry
643	JNJ BTW	2	Service industry
644	JNJ BTW	2	Service industry
645	JNJ BTW	2	Service industry
646	JNJ BTW	2	Service industry
647	JNJ BTW	2	Service industry
648	Marriott On the Move	2	Service industry
649	Marriott On the Move	2	Service industry
650	Marriott On the Move	2	Service industry
651	Marriott On the Move	2	Service industry
652	Marriott On the Move	2	Service industry
653	Marriott On the Move	2	Service industry
654	Marriott On the Move	2	Service industry
655	Marriott On the Move	2	Service industry
656	Marriott On the Move	2	Service industry
657	Marriott On the Move	2	Service industry
658	The Cleanest Line	2	Manufacturing
659	The Cleanest Line	2	Manufacturing
660	The Cleanest Line	2	Manufacturing
661	Hu Yoshida	2	Manufacturing
662	Hu Yoshida	2	Manufacturing
663	Hu Yoshida	2	Manufacturing
664	Hu Yoshida	2	Manufacturing
665	Hu Yoshida	2	Manufacturing
666	Hu Yoshida	2	Manufacturing
667	Hu Yoshida	2	Manufacturing

668	Garmin Blog	2	Manufacturing
669	Garmin Blog	2	Manufacturing
670	Garmin Blog	2	Manufacturing
671	Garmin Blog	2	Manufacturing
672	Garmin Blog	2	Manufacturing
673	Garmin Blog	5	Manufacturing
674	Garmin Blog	2	Manufacturing
675	Garmin Blog	2	Manufacturing
676	Garmin Blog	5	Manufacturing
677	Garmin Blog	2	Manufacturing
678	Garmin Blog	2	Manufacturing
679	Garmin Blog	2	Manufacturing
680	Garmin Blog	2	Manufacturing
681	Garmin Blog	2	Manufacturing
682	Garmin Blog	4	Manufacturing
683	Garmin Blog	5	Manufacturing
684	Garmin Blog	1	Manufacturing
685	Garmin Blog	2	Manufacturing
686	Garmin Blog	2	Manufacturing
687	Garmin Blog	2	Manufacturing
688	Infusionsoft Blog	2	Information and communication
689	Hu Yoshida	2	Manufacturing
690	Hu Yoshida	2	Manufacturing
691	Hu Yoshida	5	Manufacturing
692	Hu Yoshida	1	Manufacturing
693	Hu Yoshida	2	Manufacturing
694	Hu Yoshida	2	Manufacturing
695	Hu Yoshida	2	Manufacturing
696	Hu Yoshida	2	Manufacturing
697	Hu Yoshida	2	Manufacturing
698	Hu Yoshida	2	Manufacturing
699	Hu Yoshida	4	Manufacturing
700	Hu Yoshida	2	Manufacturing
701	Hu Yoshida	2	Manufacturing
702	Hu Yoshida	2	Manufacturing
703	Hu Yoshida	2	Manufacturing
704	Hu Yoshida	1	Manufacturing
705	Hu Yoshida	2	Manufacturing
706	Hu Yoshida	2	Manufacturing
707	Hu Yoshida	4	Manufacturing
708	Hu Yoshida	2	Manufacturing
709	Hu Yoshida	5	Manufacturing
710	Hu Yoshida	2	Manufacturing

711	Hu Yoshida	2	Manufacturing
712	Hu Yoshida	2	Manufacturing
713	Hu Yoshida	2	Manufacturing
714	Hu Yoshida	2	Manufacturing
715	Hu Yoshida	2	Manufacturing
716	Red Cross Chat	2	Activities of extraterritorial organizations and bodies
717	Red Cross Chat	2	Activities of extraterritorial organizations and bodies
718	Red Cross Chat	2	Activities of extraterritorial organizations and bodies
719	Red Cross Chat	2	Activities of extraterritorial organizations and bodies
720		.	
721	Red Cross Chat	2	Activities of extraterritorial organizations and bodies
722		.	
723	Red Cross Chat	5	Activities of extraterritorial organizations and bodies
724	Red Cross Chat	2	Activities of extraterritorial organizations and bodies
725	Red Cross Chat	2	Activities of extraterritorial organizations and bodies
726	Toyota Open Road Blog	2	Manufacturing
727	Toyota Open Road Blog	2	Manufacturing
728	Toyota Open Road Blog	2	Manufacturing
729	Toyota Open Road Blog	2	Manufacturing
730	Toyota Open Road Blog	1	Manufacturing
731	Toyota Open Road Blog	1	Manufacturing
732	Toyota Open Road Blog	2	Manufacturing
733	Toyota Open Road Blog	4	Manufacturing
734	Toyota Open Road Blog	2	Manufacturing
735	Toyota Open Road Blog	2	Manufacturing
736	Logos Bible Software Blog	5	Information Technology
737	Logos Bible Software Blog	5	Information Technology
738	Logos Bible Software Blog	2	Information Technology
739	Logos Bible Software Blog	5	Information Technology
740	Logos Bible Software Blog	2	Information Technology
741	Logos Bible Software Blog	2	Information Technology
742	Logos Bible Software Blog	2	Information Technology
743	Logos Bible Software Blog	2	Information Technology
744	Logos Bible Software Blog	2	Information Technology
745	Logos Bible Software Blog	2	Information Technology

746	Logos Bible Software Blog	5	Information Technology
747	Logos Bible Software Blog	4	Information Technology
748	Logos Bible Software Blog	2	Information Technology
749	Logos Bible Software Blog	4	Information Technology
750	Logos Bible Software Blog	4	Information Technology
751	Logos Bible Software Blog	2	Information Technology
752	Logos Bible Software Blog	5	Information Technology
753	Logos Bible Software Blog	4	Information Technology
754	Logos Bible Software Blog	2	Information Technology
755	Logos Bible Software Blog	2	Information Technology
756	Logos Bible Software Blog	2	Information Technology
757	Logos Bible Software Blog	2	Information Technology
758	Logos Bible Software Blog	2	Information Technology
759	Logos Bible Software Blog	2	Information Technology
760	Logos Bible Software Blog	5	Information Technology
761	Logos Bible Software Blog	4	Information Technology
762	Logos Bible Software Blog	2	Information Technology
763	Logos Bible Software Blog	2	Information Technology
764	Logos Bible Software Blog	2	Information Technology
765	Logos Bible Software Blog	2	Information Technology
766	Marketing Conversation	5	Information and communication
767	Marketing Conversation	5	Information and communication
768	Marketing Conversation	5	Information and communication
769	Marketing Conversation	2	Information and communication
770	Marketing Conversation	2	Information and communication
771	Marketing Conversation	2	Information and communication
772	Marketing Conversation	2	Information and communication
773	Marketing Conversation	5	Information and communication
774	Marketing Conversation	5	Information and communication
775	Marketing Conversation	5	Information and communication
776	Marketing Conversation	2	Information and communication
777	Marketing Conversation	5	Information and communication
778	Marketing Conversation	4	Information and communication
779	Marketing Conversation	2	Information and communication
780	Marketing Conversation	4	Information and communication
781	Marketing Conversation	4	Information and communication
782	Marketing Conversation	4	Information and communication
783	Marketing Conversation	5	Information and communication
784	Marketing Conversation	5	Information and communication
785	Marketing Conversation	5	Information and communication
786	Marketing Conversation	5	Information and communication
787	Oracle Blogs	.	
788	Oracle Blogs	2	Information Technology

789	Oracle Blogs	2	Information Technology
790	Oracle Blogs	.	
791	Oracle Blogs	2	Information Technology
792	Oracle Blogs	2	Information Technology
793	Oracle Blogs	2	Information Technology
794	Oracle Blogs	1	Information Technology
795	Oracle Blogs	2	Information Technology
796	Oracle Blogs	2	Information Technology
797	Oracle Blogs	2	Information Technology
798	Oracle Blogs	.	
799	Oracle Blogs	2	Information Technology
800	Oracle Blogs	2	Information Technology
801	Oracle Blogs	5	Information Technology
802	Oracle Blogs	.	
803	Oracle Blogs	.	
804	Oracle Blogs	2	Information Technology
805	Oracle Blogs	2	Information Technology
806	Oracle Blogs	2	Information Technology
807	Oracle Blogs	5	Information Technology
808	Oracle Blogs	5	Information Technology
809	Oracle Blogs	2	Information Technology
810	Oracle Blogs	2	Information Technology
811	Oracle Blogs	2	Information Technology
812	Oracle Blogs	.	
813	Oracle Blogs	2	Information Technology
814	Oracle Blogs	2	Information Technology
815	Rubbermaid - Adventures in Organization	5	Manufacturing
816	Rubbermaid - Adventures in Organization	2	Manufacturing
817	Rubbermaid - Adventures in Organization	5	Manufacturing
818	Rubbermaid - Adventures in Organization	5	Manufacturing
819	Rubbermaid - Adventures in Organization	5	Manufacturing
820	Rubbermaid - Adventures in Organization	4	Manufacturing
821	Rubbermaid - Adventures in Organization	5	Manufacturing
822	Rubbermaid - Adventures in Organization	5	Manufacturing
823	Rubbermaid - Adventures in	5	Manufacturing

	Organization		
824	Rubbermaid - Adventures in Organization	5	Manufacturing
825	Rubbermaid - Adventures in Organization	5	Manufacturing
826	Rubbermaid - Adventures in Organization	2	Manufacturing
827	Rubbermaid - Adventures in Organization	5	Manufacturing
828	Rubbermaid - Adventures in Organization	2	Manufacturing
829	Rubbermaid - Adventures in Organization	5	Manufacturing
830	Rubbermaid - Adventures in Organization	4	Manufacturing
831	Rubbermaid - Adventures in Organization	5	Manufacturing
832	Rubbermaid - Adventures in Organization	5	Manufacturing
833	Rubbermaid - Adventures in Organization	5	Manufacturing
834	Rubbermaid - Adventures in Organization	5	Manufacturing
835	RS Consulting blog	2	Information and communication
836	RS Consulting blog	2	Information and communication
837	RS Consulting blog	2	Information and communication
838	RS Consulting blog	2	Information and communication
839	RS Consulting blog	5	Information and communication
840	RS Consulting blog	4	Information and communication
841	RS Consulting blog	2	Information and communication
842	RS Consulting blog	2	Information and communication
843	RS Consulting blog	2	Information and communication
844	RS Consulting blog	5	Information and communication
845	RS Consulting blog	1	Information and communication
846	RS Consulting blog	2	Information and communication
847	RS Consulting blog	2	Information and communication
848	RS Consulting blog	5	Information and communication
849	RS Consulting blog	2	Information and communication
850	RS Consulting blog	4	Information and communication
851	RS Consulting blog	2	Information and communication
852	RS Consulting blog	4	Information and communication
853	RS Consulting blog	2	Information and communication
854	RS Consulting blog	4	Information and communication

855	RS Consulting blog	2	Information and communication
856	PurinaCare Pet Insurance Blog	2	Service industry
857	PurinaCare Pet Insurance Blog	5	Service industry
858	PurinaCare Pet Insurance Blog	5	Service industry
859	PurinaCare Pet Insurance Blog	4	Service industry
860	PurinaCare Pet Insurance Blog	2	Service industry
861	PurinaCare Pet Insurance Blog	2	Service industry
862	PurinaCare Pet Insurance Blog	2	Service industry
863	PurinaCare Pet Insurance Blog	5	Service industry
864	PurinaCare Pet Insurance Blog	2	Service industry
865	PurinaCare Pet Insurance Blog	2	Service industry
866	PurinaCare Pet Insurance Blog	2	Service industry
867	PurinaCare Pet Insurance Blog	2	Service industry
868	PurinaCare Pet Insurance Blog	5	Service industry
869	PurinaCare Pet Insurance Blog	.	
870	PurinaCare Pet Insurance Blog	4	Service industry
871	PurinaCare Pet Insurance Blog	4	Service industry
872	PurinaCare Pet Insurance Blog	5	Service industry
873	PurinaCare Pet Insurance Blog	5	Service industry
874	PurinaCare Pet Insurance Blog	5	Service industry
875	PurinaCare Pet Insurance Blog	2	Service industry
876	RS Consulting blog	2	Information and communication
877	RS Consulting blog	4	Information and communication
878	RS Consulting blog	.	
879	RS Consulting blog	2	Information and communication
880	RS Consulting blog	4	Information and communication
881	RS Consulting blog	2	Information and communication
882	RS Consulting blog	2	Information and communication
883	RS Consulting blog	4	Information and communication
884	RS Consulting blog	2	Information and communication
885	RS Consulting blog	2	Information and communication
886	RS Consulting blog	2	Information and communication
887	RS Consulting blog	2	Information and communication
888	RS Consulting blog	2	Information and communication
889	RS Consulting blog	4	Information and communication
890	RS Consulting blog	2	Information and communication
891	FedEx Citizenship Blog	2	Service industry
892	FedEx Citizenship Blog	2	Service industry
893	FedEx Citizenship Blog	2	Service industry
894	FedEx Citizenship Blog	2	Service industry
895	FedEx Citizenship Blog	2	Service industry
896	FedEx Citizenship Blog	2	Service industry
897	FedEx Citizenship Blog	1	Service industry

898	FedEx Citizenship Blog	2	Service industry
899	FedEx Citizenship Blog	2	Service industry
900	FedEx Citizenship Blog	2	Service industry
901	The LinkedIn Blog	1	E-commerce and online social networking service
902	The LinkedIn Blog	2	E-commerce and online social networking service
903	The LinkedIn Blog	4	E-commerce and online social networking service
904	The LinkedIn Blog	2	E-commerce and online social networking service
905	The LinkedIn Blog	1	E-commerce and online social networking service
906	The LinkedIn Blog	2	E-commerce and online social networking service
907	The LinkedIn Blog	2	E-commerce and online social networking service
908	The LinkedIn Blog	2	E-commerce and online social networking service
909	The LinkedIn Blog	2	E-commerce and online social networking service
910	The LinkedIn Blog	2	E-commerce and online social networking service
911	The LinkedIn Blog	2	E-commerce and online social networking service
912	The LinkedIn Blog	2	E-commerce and online social networking service
913	The LinkedIn Blog	1	E-commerce and online social networking service
914	The LinkedIn Blog	2	E-commerce and online social networking service
915	The LinkedIn Blog	2	E-commerce and online social networking service
916	The LinkedIn Blog	2	E-commerce and online social networking service
917	The LinkedIn Blog	5	E-commerce and online social networking service
918	The LinkedIn Blog	2	E-commerce and online social networking service
919	The LinkedIn Blog	2	E-commerce and online social networking service
920	The LinkedIn Blog	1	E-commerce and online social networking service

921	The LinkedIn Blog	2	E-commerce and online social networking service
922	The LinkedIn Blog	2	E-commerce and online social networking service
923	The LinkedIn Blog	2	E-commerce and online social networking service
924	The LinkedIn Blog	2	E-commerce and online social networking service
925	The LinkedIn Blog	1	E-commerce and online social networking service
926	The LinkedIn Blog	2	E-commerce and online social networking service
927	The LinkedIn Blog	4	E-commerce and online social networking service
928	The LinkedIn Blog	1	E-commerce and online social networking service
929	Tyson Foods Hunger Relief	2	Information and communication
930	Tyson Foods Hunger Relief	2	Information and communication
931	Tyson Foods Hunger Relief	2	Information and communication
932	Tyson Foods Hunger Relief	2	Information and communication
933	Tyson Foods Hunger Relief	4	Information and communication
934	Tyson Foods Hunger Relief	2	Information and communication
935	Tyson Foods Hunger Relief	2	Information and communication
936	Tyson Foods Hunger Relief	1	Information and communication
937	Tyson Foods Hunger Relief	4	Information and communication
938	Tyson Foods Hunger Relief	2	Information and communication
939	Tyson Foods Hunger Relief	1	Information and communication
940	Tyson Foods Hunger Relief	2	Information and communication
941	Tyson Foods Hunger Relief	2	Information and communication
942	Tyson Foods Hunger Relief	.	
943	Tyson Foods Hunger Relief	.	
944	Tyson Foods Hunger Relief	.	
945	Tyson Foods Hunger Relief	.	
946	Tyson Foods Hunger Relief	.	
947	Tyson Foods Hunger Relief	2	Information and communication
948	Tyson Foods Hunger Relief	2	Information and communication
949	Tyson Foods Hunger Relief	2	Information and communication
950	Tyson Foods Hunger Relief	2	Information and communication
951	Tyson Foods Hunger Relief	2	Information and communication
952	Tyson Foods Hunger Relief	2	Information and communication
953	Tyson Foods Hunger Relief	1	Information and communication
954	The Unboxed Thoughts Blog	4	Information and communication
955	The Unboxed Thoughts Blog	4	Information and communication

956	The Unboxed Thoughts Blog	1	Information and communication
957	The Unboxed Thoughts Blog	2	Information and communication
958	The Unboxed Thoughts Blog	2	Information and communication
959	The Unboxed Thoughts Blog	2	Information and communication
960	The Unboxed Thoughts Blog	2	Information and communication
961	The Unboxed Thoughts Blog	1	Information and communication
962	The Unboxed Thoughts Blog	2	Information and communication
963	The Unboxed Thoughts Blog	2	Information and communication
964	The Unboxed Thoughts Blog	2	Information and communication
965	The Unboxed Thoughts Blog	2	Information and communication
966	The Unboxed Thoughts Blog	4	Information and communication
967	The Unboxed Thoughts Blog	5	Information and communication
968	The Unboxed Thoughts Blog	4	Information and communication
969	The Unboxed Thoughts Blog	2	Information and communication
970	The Unboxed Thoughts Blog	5	Information and communication
971	The Unboxed Thoughts Blog	5	Information and communication
972	The Unboxed Thoughts Blog	4	Information and communication
973	The Unboxed Thoughts Blog	2	Information and communication
974	The Unboxed Thoughts Blog	2	Information and communication
975	The Unboxed Thoughts Blog	2	Information and communication
976	The Unboxed Thoughts Blog	2	Information and communication
977	The Unboxed Thoughts Blog	2	Information and communication
978	The Unboxed Thoughts Blog	4	Information and communication
979	Upside Learning Blog	5	Information Technology
980	Upside Learning Blog	2	Information Technology
981	Upside Learning Blog	5	Information Technology
982	Upside Learning Blog	2	Information Technology
983	Upside Learning Blog	2	Information Technology
984	Upside Learning Blog	2	Information Technology
985	Upside Learning Blog	2	Information Technology
986	Upside Learning Blog	2	Information Technology
987	Upside Learning Blog	2	Information Technology
988	Upside Learning Blog	2	Information Technology
989	Upside Learning Blog	2	Information Technology
990	Upside Learning Blog	2	Information Technology
991	Upside Learning Blog	1	Information Technology
992	Upside Learning Blog	2	Information Technology
993	Upside Learning Blog	4	Information Technology
994	Upside Learning Blog	2	Information Technology
995	Upside Learning Blog	2	Information Technology
996	Upside Learning Blog	2	Information Technology
997	Upside Learning Blog	2	Information Technology
998	Upside Learning Blog	2	Information Technology

999	Upside Learning Blog	2	Information Technology
1000	Upside Learning Blog	2	Information Technology
1001	Upside Learning Blog	2	Information Technology
1002	Upside Learning Blog	2	Information Technology
1003	Upside Learning Blog	4	Information Technology
1004	Upside Learning Blog	2	Information Technology
1005	Upside Learning Blog	2	Information Technology
1006	Upside Learning Blog	5	Information Technology
1007	Upside Learning Blog	2	Information Technology
1008	Upside Learning Blog	2	Information Technology
1009	Verizon-PolicyBlog	2	Information and communication
1010	Verizon-PolicyBlog	1	Information and communication
1011	Verizon-PolicyBlog	2	Information and communication
1012	Verizon-PolicyBlog	1	Information and communication
1013	Verizon-PolicyBlog	2	Information and communication
1014	Verizon-PolicyBlog	1	Information and communication
1015	Verizon-PolicyBlog	2	Information and communication
1016	Verizon-PolicyBlog	2	Information and communication
1017	Verizon-PolicyBlog	2	Information and communication
1018	Verizon-PolicyBlog	4	Information and communication
1019	Verizon-PolicyBlog	2	Information and communication
1020	Verizon-PolicyBlog	2	Information and communication
1021	Verizon-PolicyBlog	2	Information and communication
1022	Verizon-PolicyBlog	4	Information and communication
1023	Verizon-PolicyBlog	1	Information and communication
1024	Verizon-PolicyBlog	4	Information and communication
1025	Verizon-PolicyBlog	2	Information and communication
1026	Verizon-PolicyBlog	2	Information and communication
1027	Verizon-PolicyBlog	1	Information and communication
1028	Verizon-PolicyBlog	2	Information and communication
1029	Verizon-PolicyBlog	4	Information and communication
1030	Verizon-PolicyBlog	2	Information and communication
1031	Verizon-PolicyBlog	2	Information and communication
1032	Verizon-PolicyBlog	1	Information and communication
1033	Verizon-PolicyBlog	2	Information and communication
1034	Xerox Blogs	2	Information Technology
1035	Xerox Blogs	2	Information Technology
1036	Xerox Blogs	1	Information Technology
1037	Xerox Blogs	2	Information Technology
1038	Xerox Blogs	5	Information Technology
1039	Xerox Blogs	5	Information Technology
1040	Xerox Blogs	2	Information Technology
1041	Xerox Blogs	2	Information Technology

1042	Xerox Blogs	2	Information Technology
1043	Xerox Blogs	.	
1044	Xerox Blogs	.	
1045	Xerox Blogs	5	Information Technology
1046	Xerox Blogs	2	Information Technology
1047	Xerox Blogs	2	Information Technology
1048	Xerox Blogs	2	Information Technology
1049	Xerox Blogs	2	Information Technology
1050	Xerox Blogs	1	Information Technology
1051	Xerox Blogs	5	Information Technology
1052	Xerox Blogs	2	Information Technology
1053	Xerox Blogs	2	Information Technology
1054	Xerox Blogs	1	Information Technology
1055	Xerox Blogs	1	Information Technology
1056	Xerox Blogs	2	Information Technology
1057	Xerox Blogs	5	Information Technology
1058	Xerox Blogs	1	Information Technology